



(12) BẢN MÔ TẢ SÁNG CHẾ THUỘC BẰNG ĐỘC QUYỀN SÁNG CHẾ

(19) Cộng hòa xã hội chủ nghĩa Việt Nam (VN)
CỤC SỞ HỮU TRÍ TUỆ

(11)



1-0024824

(51)⁷ G10L 25/78

(13) B

(21) 1-2016-03722

(22) 01/12/2014

(86) PCT/CN2014/092694 01/12/2014

(87) WO2015/135344 17/09/2015

(30) 201410090386.X 12/03/2014 CN

(45) 25/08/2020 389

(43) 26/12/2016 345A

(73) HUAWEI TECHNOLOGIES CO., LTD. (CN)

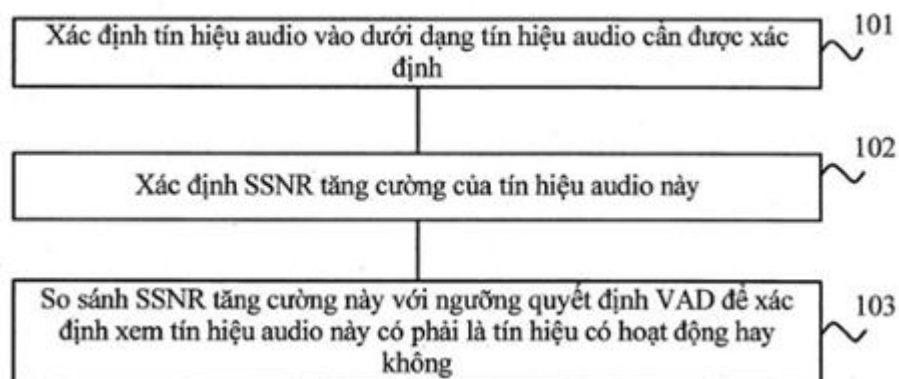
Huawei Administration Building, Bantian, Longgang, Shenzhen, Guangdong 518129, China

(72) WANG, Zhe (CN).

(74) Công ty Luật TNHH Phạm và Liên danh (PHAM & ASSOCIATES)

(54) PHƯƠNG PHÁP VÀ THIẾT BỊ PHÁT HIỆN TÍN HIỆU AUDIO

(57) Sáng chế đề cập đến phương pháp và thiết bị phát hiện tín hiệu audio, trong đó phương pháp này bao gồm các bước: xác định tín hiệu audio vào dưới dạng tín hiệu audio cần được xác định; xác định tỉ số tín hiệu trên tạp âm theo phân đoạn (Segmental Signal-to-Noise Ratio - SSNR) tăng cường của tín hiệu audio, trong đó SSNR tăng cường này lớn hơn SSNR tham chiếu; và so sánh SSNR tăng cường này với ngưỡng quyết định phát hiện hoạt động thoại (Voice Activity Detection - VAD) để xác định xem tín hiệu audio này có phải là tín hiệu có hoạt động hay không. Theo phương pháp và thiết bị theo sáng chế, âm thoại có hoạt động và âm thoại không có hoạt động có thể được phân biệt một cách chính xác.



Lĩnh vực kỹ thuật được đề cập

Sáng chế liên quan đến các công nghệ xử lý tín hiệu, cụ thể là đề cập đến phương pháp và thiết bị phát hiện tín hiệu audio.

Tình trạng kỹ thuật của sáng chế

Phát hiện hoạt động thoại (Voice Activity Detection - VAD) là công nghệ chủ chốt được sử dụng rộng rãi trong các lĩnh vực như truyền thông thoại và tương tác người-máy. VAD còn có thể được gọi là phát hiện hoạt động âm thanh (Sound Activity Detection - SAD). VAD được dùng để phát hiện xem có tín hiệu có hoạt động nào trong tín hiệu audio vào hay không, trong đó tín hiệu có hoạt động là tương đối so với tín hiệu không hoạt động (chẳng hạn tạp âm nền của môi trường và âm thoại câm). Các tín hiệu có hoạt động thông thường bao gồm tiếng nói, âm nhạc, v.v.. Nguyên lý của VAD là một hoặc nhiều thông số đặc trưng được trích xuất từ tín hiệu audio vào, một hoặc nhiều giá trị đặc trưng được xác định theo một hoặc nhiều thông số đặc trưng này, và sau đó một hoặc nhiều giá trị đặc trưng này được so sánh với một hoặc nhiều ngưỡng.

Theo giải pháp đã biết, phương pháp phát hiện tín hiệu có hoạt động dựa trên tỉ số tín hiệu trên tạp âm theo phân đoạn (Segmental Signal-to-Noise Ratio - SSNR) bao gồm các bước: chia tín hiệu audio vào thành các tín hiệu dải con trên dải tần, tính toán năng lượng của tín hiệu audio này trên mỗi dải con, và so sánh năng lượng của tín hiệu audio trên mỗi dải con với năng lượng ước lượng được của tín hiệu tạp âm nền trên mỗi dải con, để thu được tỉ số tín hiệu trên tạp âm (Signal-to-Noise Ratio - SNR) của tín hiệu audio trên mỗi dải con; và sau đó xác định SSNR theo SNR dải con của mỗi dải con, và so sánh SSNR này với ngưỡng quyết định

VAD định trước, trong đó nếu SSNR này lớn hơn ngưỡng quyết định VAD này, thì tín hiệu audio này là tín hiệu có hoạt động, còn nếu SSNR này không lớn hơn ngưỡng quyết định VAD này thì tín hiệu audio này là tín hiệu không có hoạt động.

Phương pháp thông thường để tính SSNR là cộng tất cả các SNR dải con của tín hiệu audio với nhau, và kết quả thu được là SSNR. Ví dụ, SSNR có thể được xác định bằng công thức 1.1:

$$SSNR = \sum_{k=0}^{N-1} snr(k)$$

Công thức 1.1

trong đó k biểu thị dải con thứ k, snr(k) biểu thị SNR dải con của dải con thứ k, và N biểu thị tổng số dải con mà tín hiệu audio được chia thành.

Khi phương pháp tính SSNR nêu trên được dùng để phát hiện âm thoại có hoạt động, thì có thể xảy ra trường hợp không phát hiện được âm thoại có hoạt động.

Bản chất kỹ thuật của sáng chế

Mục đích của sáng chế là đề xuất phương pháp và thiết bị phát hiện tín hiệu audio, để có thể phân biệt chính xác giữa âm thoại có hoạt động và âm thoại không có hoạt động.

Theo khía cạnh thứ nhất, theo một phương án, sáng chế đề xuất phương pháp phát hiện tín hiệu audio, trong đó phương pháp này bao gồm các bước: xác định tín hiệu audio vào dưới dạng tín hiệu audio cần được xác định; xác định tỉ số tín hiệu trên tạp âm theo phân đoạn (Segmental Signal-to-Noise Ratio - SSNR) tăng cường của tín hiệu audio, trong đó SSNR tăng cường này lớn hơn SSNR tham chiếu; và so sánh SSNR tăng cường này với ngưỡng quyết định phát hiện hoạt động thoại (Voice Activity Detection - VAD) để xác định xem tín hiệu audio này có phải là tín hiệu có hoạt động hay không.

Dựa vào khía cạnh thứ nhất, theo cách thức thực hiện khả thi thứ nhất

của khía cạnh thứ nhất, bước xác định tín hiệu audio vào dưới dạng tín hiệu audio cần được xác định là bước: xác định tín hiệu audio này dưới dạng tín hiệu audio cần được xác định theo tỉ số tín hiệu trên tạp âm (Signal-to-Noise Ratio - SNR) theo dải con của tín hiệu audio này.

Dựa vào cách thức thực hiện khả thi thứ nhất của khía cạnh thứ nhất, theo cách thức thực hiện khả thi thứ hai của khía cạnh thứ nhất, bước xác định tín hiệu audio vào dưới dạng tín hiệu audio cần được xác định là bước: xác định tín hiệu audio này dưới dạng tín hiệu audio cần được xác định trong trường hợp số lượng dải con đầu cao tần, mà nằm trong tín hiệu audio này và có các SNR dải con lớn hơn ngưỡng định trước thứ nhất, là lớn hơn số lượng thứ nhất.

Dựa vào cách thức thực hiện khả thi thứ nhất của khía cạnh thứ nhất, theo cách thức thực hiện khả thi thứ ba của khía cạnh thứ nhất, bước xác định tín hiệu audio vào dưới dạng tín hiệu audio cần được xác định là bước: xác định tín hiệu audio này dưới dạng tín hiệu audio cần được xác định trong trường hợp số lượng dải con đầu cao tần, mà nằm trong tín hiệu audio này và có các SNR dải con lớn hơn ngưỡng định trước thứ nhất, là lớn hơn số lượng thứ hai, và số lượng dải con đầu hạ tần, mà nằm trong tín hiệu audio này và có các SNR dải con nhỏ hơn ngưỡng định trước thứ hai, là lớn hơn số lượng thứ ba.

Dựa vào cách thức thực hiện khả thi thứ nhất của khía cạnh thứ nhất, theo cách thức thực hiện khả thi thứ tư của khía cạnh thứ nhất, bước xác định tín hiệu audio vào dưới dạng tín hiệu audio cần được xác định là bước: xác định tín hiệu audio này dưới dạng tín hiệu audio cần được xác định trong trường hợp số lượng dải con, mà nằm trong tín hiệu audio này và có giá trị của các SNR dải con lớn hơn ngưỡng định trước thứ ba, là lớn hơn số lượng thứ tư.

Dựa vào khía cạnh thứ nhất, theo cách thức thực hiện khả thi thứ năm của khía cạnh thứ nhất, bước xác định tín hiệu audio vào dưới dạng tín hiệu

audio cần được xác định là bước: xác định tín hiệu audio này dưới dạng tín hiệu audio cần được xác định trong trường hợp xác định được rằng tín hiệu audio này là tín hiệu âm vô thanh.

Dựa vào cách thức thực hiện khả thi thứ hai hoặc cách thức thực hiện khả thi thứ ba của khía cạnh thứ nhất, theo cách thức thực hiện khả thi thứ sáu của khía cạnh thứ nhất, bước xác định SSNR tăng cường của tín hiệu audio nêu trên là các bước: xác định trọng số của SNR dải con của mỗi dải con trong tín hiệu audio này, trong đó trọng số của SNR dải con của dải con đầu cao tần mà có SNR dải con lớn hơn ngưỡng định trước thứ nhất là lớn hơn trọng số của SNR dải con của dải con khác; và xác định SSNR tăng cường theo SNR dải con của mỗi dải con và trọng số của SNR dải con của mỗi dải con trong tín hiệu audio này.

Dựa vào khía cạnh thứ nhất hoặc cách thức thực hiện khả thi bất kì của cách thức thực hiện khả thi thứ nhất của khía cạnh thứ nhất đến cách thức thực hiện khả thi thứ năm của khía cạnh thứ nhất, theo cách thức thực hiện khả thi thứ bảy của khía cạnh thứ nhất, bước xác định SSNR tăng cường của tín hiệu audio nêu trên là các bước: xác định SSNR tham chiếu của tín hiệu audio này; và xác định SSNR tăng cường theo SSNR tham chiếu của tín hiệu audio này.

Dựa vào cách thức thực hiện khả thi thứ bảy của khía cạnh thứ nhất, theo cách thức thực hiện khả thi thứ tám của khía cạnh thứ nhất, bước xác định SSNR tăng cường theo SSNR tham chiếu của tín hiệu audio là bước: xác định SSNR tăng cường bằng công thức sau: $SSNR' = x * SSNR + y$, trong đó SSNR biểu thị SSNR tham chiếu, SSNR' biểu thị SSNR tăng cường, và x và y biểu thị các thông số tăng cường.

Dựa vào cách thức thực hiện khả thi thứ bảy của khía cạnh thứ nhất, theo cách thức thực hiện khả thi thứ chín của khía cạnh thứ nhất, bước xác định SSNR tăng cường theo SSNR tham chiếu của tín hiệu audio là bước:

xác định SSNR tăng cường bằng công thức sau:
 $SSNR' = f(x) * SSNR + h(y)$, trong đó SSNR biểu thị SSNR tham chiếu, SSNR' biểu thị SSNR tăng cường, và $f(x)$ và $h(y)$ biểu thị các hàm tăng cường.

Dựa vào khía cạnh thứ nhất hoặc bất kì trong số những cách thức thực hiện khả thi nêu trên của khía cạnh thứ nhất, theo cách thức thực hiện khả thi thứ mười của khía cạnh thứ nhất, trước bước so sánh SSNR tăng cường với ngưỡng quyết định VAD, phương pháp này còn bao gồm bước: dùng thuật toán định trước để giảm ngưỡng quyết định VAD, để thu được ngưỡng quyết định VAD đã giảm; và bước so sánh SSNR tăng cường với ngưỡng quyết định VAD để xác định xem tín hiệu audio có phải là tín hiệu có hoạt động hay không cụ thể là bước: so sánh SSNR tăng cường với ngưỡng quyết định VAD đã giảm để xác định xem tín hiệu audio này có phải là tín hiệu có hoạt động hay không.

Theo khía cạnh thứ hai, theo một phương án, sáng chế đề xuất phương pháp phát hiện tín hiệu audio, trong đó phương pháp này bao gồm các bước: xác định tín hiệu audio vào dưới dạng tín hiệu audio cần được xác định; xác định trọng số của SNR dải con của mỗi dải con trong tín hiệu audio này, trong đó trọng số của SNR dải con của dải con đầu cao tần mà có SNR dải con lớn hơn ngưỡng định trước thứ nhất là lớn hơn trọng số của SNR dải con của dải con khác; xác định SSNR tăng cường theo SNR dải con của mỗi dải con và trọng số của SNR dải con của mỗi dải con trong tín hiệu audio này, trong đó SSNR tăng cường này là lớn hơn SSNR tham chiếu; và so sánh SSNR tăng cường này với ngưỡng quyết định VAD để xác định xem tín hiệu audio này có phải là tín hiệu có hoạt động hay không.

Dựa vào khía cạnh thứ hai, theo cách thức thực hiện khả thi thứ nhất của khía cạnh thứ hai, bước xác định tín hiệu audio vào dưới dạng tín hiệu audio cần được xác định là bước: xác định tín hiệu audio này dưới dạng tín

hiệu audio cần được xác định theo SNR dải con của tín hiệu audio này.

Dựa vào cách thức thực hiện khả thi thứ nhất của khía cạnh thứ hai, theo cách thức thực hiện khả thi thứ hai của khía cạnh thứ hai, bước xác định tín hiệu audio vào dưới dạng tín hiệu audio cần được xác định là bước: xác định tín hiệu audio này dưới dạng tín hiệu audio cần được xác định trong trường hợp số lượng dải con đầu cao tần, mà nằm trong tín hiệu audio này và có các SNR dải con lớn hơn ngưỡng định trước thứ nhất, là lớn hơn số lượng thứ nhất.

Dựa vào cách thức thực hiện khả thi thứ nhất của khía cạnh thứ hai, theo cách thức thực hiện khả thi thứ ba của khía cạnh thứ hai, bước xác định tín hiệu audio vào dưới dạng tín hiệu audio cần được xác định là bước: xác định tín hiệu audio này dưới dạng tín hiệu audio cần được xác định trong trường hợp số lượng dải con đầu cao tần, mà nằm trong tín hiệu audio này và có các SNR dải con lớn hơn ngưỡng định trước thứ nhất, là lớn hơn số lượng thứ hai, và số lượng dải con đầu hạ tần, mà nằm trong tín hiệu audio này và có các SNR dải con nhỏ hơn ngưỡng định trước thứ hai, là lớn hơn số lượng thứ ba.

Theo khía cạnh thứ ba, theo một phương án, sáng chế đề xuất phương pháp phát hiện tín hiệu audio, trong đó phương pháp này bao gồm các bước: xác định tín hiệu audio vào dưới dạng tín hiệu audio cần được xác định; thu thập SSNR tham chiếu của tín hiệu audio này; dùng thuật toán định trước để giảm ngưỡng quyết định VAD tham chiếu, để thu được ngưỡng quyết định VAD đã giảm; và so sánh SSNR tham chiếu với ngưỡng quyết định VAD đã giảm để xác định xem tín hiệu audio này có phải là tín hiệu có hoạt động hay không.

Dựa vào khía cạnh thứ ba, theo cách thức thực hiện khả thi thứ nhất của khía cạnh thứ ba, bước xác định tín hiệu audio vào dưới dạng tín hiệu audio cần được xác định là bước: xác định tín hiệu audio này dưới dạng tín hiệu audio cần được xác định theo SNR dải con của tín hiệu audio này.

Dựa vào cách thức thực hiện khả thi thứ nhất của khía cạnh thứ ba, theo cách thức thực hiện khả thi thứ hai của khía cạnh thứ ba, bước xác định tín hiệu audio vào dưới dạng tín hiệu audio cần được xác định là bước: xác định tín hiệu audio này dưới dạng tín hiệu audio cần được xác định trong trường hợp số lượng dải con đầu cao tần, mà nằm trong tín hiệu audio này và có các SNR dải con lớn hơn ngưỡng định trước thứ nhất, là lớn hơn số lượng thứ nhất.

Dựa vào cách thức thực hiện khả thi thứ nhất của khía cạnh thứ ba, theo cách thức thực hiện khả thi thứ ba của khía cạnh thứ ba, bước xác định tín hiệu audio vào dưới dạng tín hiệu audio cần được xác định là bước: xác định tín hiệu audio này dưới dạng tín hiệu audio cần được xác định trong trường hợp số lượng dải con đầu cao tần, mà nằm trong tín hiệu audio này và có các SNR dải con lớn hơn ngưỡng định trước thứ nhất, là lớn hơn số lượng thứ hai, và số lượng dải con đầu hạ tần, mà nằm trong tín hiệu audio này và có các SNR dải con nhỏ hơn ngưỡng định trước thứ hai, là lớn hơn số lượng thứ ba.

Dựa vào cách thức thực hiện khả thi thứ nhất của khía cạnh thứ ba, theo cách thức thực hiện khả thi thứ tư của khía cạnh thứ ba, bước xác định tín hiệu audio vào dưới dạng tín hiệu audio cần được xác định là bước: xác định tín hiệu audio này dưới dạng tín hiệu audio cần được xác định trong trường hợp số lượng dải con, mà nằm trong tín hiệu audio này và có giá trị của các SNR dải con lớn hơn ngưỡng định trước thứ ba, là lớn hơn số lượng thứ tư.

Dựa vào khía cạnh thứ ba, theo cách thức thực hiện khả thi thứ năm của khía cạnh thứ ba, bước xác định tín hiệu audio vào dưới dạng tín hiệu audio cần được xác định là bước: xác định tín hiệu audio này dưới dạng tín hiệu audio cần được xác định trong trường hợp xác định được rằng tín hiệu audio này là tín hiệu âm vô thanh.

Theo khía cạnh thứ tư, theo một phương án, sáng chế đề xuất thiết bị, trong đó thiết bị này bao gồm: khối xác định thứ nhất, được tạo cấu hình để

xác định tín hiệu audio vào dưới dạng tín hiệu audio cần được xác định; khối xác định thứ hai, được tạo cấu hình để xác định SSNR tăng cường của tín hiệu audio này, trong đó SSNR tăng cường này lớn hơn SSNR tham chiếu; và khối xác định thứ ba, được tạo cấu hình để so sánh SSNR tăng cường này với ngưỡng quyết định VAD để xác định xem tín hiệu audio này có phải là tín hiệu có hoạt động hay không.

Dựa vào khía cạnh thứ tư, theo cách thức thực hiện khả thi thứ nhất của khía cạnh thứ tư, khối xác định thứ nhất được tạo cấu hình cụ thể để xác định tín hiệu audio này dưới dạng tín hiệu audio cần được xác định theo SNR dải con của tín hiệu audio này.

Dựa vào cách thức thực hiện khả thi thứ nhất của khía cạnh thứ tư, theo cách thức thực hiện khả thi thứ hai của khía cạnh thứ tư, khối xác định thứ nhất được tạo cấu hình cụ thể để xác định tín hiệu audio này dưới dạng tín hiệu audio cần được xác định trong trường hợp số lượng dải con đầu cao tần, mà nằm trong tín hiệu audio này và có các SNR dải con lớn hơn ngưỡng định trước thứ nhất, là lớn hơn số lượng thứ nhất.

Dựa vào cách thức thực hiện khả thi thứ nhất của khía cạnh thứ tư, theo cách thức thực hiện khả thi thứ ba của khía cạnh thứ tư, khối xác định thứ nhất được tạo cấu hình cụ thể để xác định tín hiệu audio này dưới dạng tín hiệu audio cần được xác định trong trường hợp số lượng dải con đầu cao tần, mà nằm trong tín hiệu audio này và có các SNR dải con lớn hơn ngưỡng định trước thứ nhất, là lớn hơn số lượng thứ hai, và số lượng dải con đầu hạ tần, mà nằm trong tín hiệu audio này và có các SNR dải con nhỏ hơn ngưỡng định trước thứ hai, là lớn hơn số lượng thứ ba.

Dựa vào cách thức thực hiện khả thi thứ nhất của khía cạnh thứ tư, theo cách thức thực hiện khả thi thứ tư của khía cạnh thứ tư, khối xác định thứ nhất được tạo cấu hình cụ thể để xác định tín hiệu audio này dưới dạng tín hiệu audio cần được xác định trong trường hợp số lượng dải con, mà nằm trong tín hiệu audio này và có giá trị của các SNR dải con lớn hơn ngưỡng

định trước thứ ba, là lớn hơn số lượng thứ tư.

Dựa vào khía cạnh thứ tư, theo cách thức thực hiện khả thi thứ năm của khía cạnh thứ tư, khối xác định thứ nhất được tạo cấu hình cụ thể để xác định tín hiệu audio này dưới dạng tín hiệu audio cần được xác định trong trường hợp xác định được rằng tín hiệu audio này là tín hiệu âm vô thanh.

Dựa vào cách thức thực hiện khả thi thứ hai của khía cạnh thứ tư hoặc cách thức thực hiện khả thi thứ ba của khía cạnh thứ tư, theo cách thức thực hiện khả thi thứ sáu của khía cạnh thứ tư, khối xác định thứ hai được tạo cấu hình cụ thể để xác định trọng số của SNR dải con của mỗi dải con trong tín hiệu audio này, trong đó trọng số của SNR dải con của dải con đầu cao tần mà có SNR dải con lớn hơn ngưỡng định trước thứ nhất là lớn hơn trọng số của SNR dải con của dải con khác; và xác định SSNR tăng cường theo SNR dải con của mỗi dải con và trọng số của SNR dải con của mỗi dải con trong tín hiệu audio này.

Dựa vào khía cạnh thứ tư hoặc cách thức thực hiện khả thi bất kì của cách thức thực hiện khả thi thứ nhất của khía cạnh thứ tư đến cách thức thực hiện khả thi thứ năm của khía cạnh thứ tư, theo cách thức thực hiện khả thi thứ bảy của khía cạnh thứ tư, khối xác định thứ hai được tạo cấu hình cụ thể để xác định SSNR tham chiếu của tín hiệu audio này; và xác định SSNR tăng cường theo SSNR tham chiếu của tín hiệu audio này.

Dựa vào cách thức thực hiện khả thi thứ bảy của khía cạnh thứ tư, theo cách thức thực hiện khả thi thứ tám của khía cạnh thứ tư, khối xác định thứ hai được tạo cấu hình cụ thể để xác định SSNR tăng cường bằng công thức sau: $SSNR' = x * SSNR + y$, trong đó SSNR biểu thị SSNR tham chiếu, SSNR' biểu thị SSNR tăng cường, và x và y biểu thị các thông số tăng cường.

Dựa vào cách thức thực hiện khả thi thứ bảy của khía cạnh thứ tư, theo cách thức thực hiện khả thi thứ chín của khía cạnh thứ tư, khối xác định thứ

hai được tạo cấu hình cụ thể để xác định SSNR tăng cường bằng công thức sau: $SSNR' = f(x) * SSNR + h(y)$, trong đó SSNR biểu thị SSNR tham chiếu, SSNR' biểu thị SSNR tăng cường, và $f(x)$ và $h(y)$ biểu thị các hàm tăng cường.

Dựa vào khía cạnh thứ tư hoặc bất kì trong số những cách thức thực hiện khả thi nêu trên của khía cạnh thứ tư, theo cách thức thực hiện khả thi thứ mười của khía cạnh thứ tư, thiết bị này còn bao gồm khối xác định thứ tư, trong đó khối xác định thứ tư này được tạo cấu hình để dùng thuật toán định trước để giảm ngưỡng quyết định VAD, để thu được ngưỡng quyết định VAD đã giảm; và khối xác định thứ ba được tạo cấu hình cụ thể để so sánh SSNR tăng cường với ngưỡng quyết định VAD đã giảm để xác định xem tín hiệu audio này có phải là tín hiệu có hoạt động hay không.

Theo khía cạnh thứ năm, theo một phương án, sáng chế đề xuất thiết bị, trong đó thiết bị này bao gồm: khối xác định thứ nhất, được tạo cấu hình để xác định tín hiệu audio vào dưới dạng tín hiệu audio cần được xác định; khối xác định thứ hai, được tạo cấu hình để xác định trọng số của SNR dải con của mỗi dải con trong tín hiệu audio này, trong đó trọng số của SNR dải con của dải con đầu cao tần, mà có SNR dải con lớn hơn ngưỡng định trước thứ nhất, là lớn hơn trọng số của SNR dải con của dải con khác, và xác định SSNR tăng cường theo SNR dải con của mỗi dải con và trọng số của SNR dải con của mỗi dải con trong tín hiệu audio này, trong đó SSNR tăng cường này là lớn hơn SSNR tham chiếu; và khối xác định thứ ba, được tạo cấu hình để so sánh SSNR tăng cường này với ngưỡng quyết định VAD để xác định xem tín hiệu audio này có phải là tín hiệu có hoạt động hay không.

Dựa vào khía cạnh thứ năm, theo cách thức thực hiện khả thi thứ nhất của khía cạnh thứ năm, khối xác định thứ nhất được tạo cấu hình cụ thể để xác định tín hiệu audio này dưới dạng tín hiệu audio cần được xác định theo SNR dải con của tín hiệu audio này.

Dựa vào cách thức thực hiện khả thi thứ nhất của khía cạnh thứ năm, theo cách thức thực hiện khả thi thứ hai của khía cạnh thứ năm, khối xác định thứ nhất được tạo cấu hình cụ thể để xác định tín hiệu audio này dưới dạng tín hiệu audio cần được xác định trong trường hợp số lượng dải con đầu cao tần, mà nằm trong tín hiệu audio này và có các SNR dải con lớn hơn ngưỡng định trước thứ nhất, là lớn hơn số lượng thứ nhất.

Dựa vào cách thức thực hiện khả thi thứ nhất của khía cạnh thứ năm, theo cách thức thực hiện khả thi thứ ba của khía cạnh thứ năm, khối xác định thứ nhất được tạo cấu hình cụ thể để xác định tín hiệu audio này dưới dạng tín hiệu audio cần được xác định trong trường hợp số lượng dải con đầu cao tần, mà nằm trong tín hiệu audio này và có các SNR dải con lớn hơn ngưỡng định trước thứ nhất, là lớn hơn số lượng thứ hai, và số lượng dải con đầu hạ tần, mà nằm trong tín hiệu audio này và có các SNR dải con nhỏ hơn ngưỡng định trước thứ hai, là lớn hơn số lượng thứ ba.

Theo khía cạnh thứ sáu, theo một phương án, sáng chế đề xuất thiết bị, trong đó thiết bị này bao gồm: khối xác định thứ nhất, được tạo cấu hình để xác định tín hiệu audio vào dưới dạng tín hiệu audio cần được xác định; khối xác định thứ hai, được tạo cấu hình để thu thập SSNR tham chiếu của tín hiệu audio này; khối xác định thứ ba, được tạo cấu hình để dùng thuật toán định trước để giảm ngưỡng quyết định VAD tham chiếu, để thu được ngưỡng quyết định VAD đã giảm; và khối xác định thứ tư, được tạo cấu hình để so sánh SSNR tham chiếu với ngưỡng quyết định VAD đã giảm để xác định xem tín hiệu audio này có phải là tín hiệu có hoạt động hay không.

Dựa vào khía cạnh thứ sáu, theo cách thức thực hiện khả thi thứ nhất của khía cạnh thứ sáu, khối xác định thứ nhất được tạo cấu hình cụ thể để xác định tín hiệu audio này dưới dạng tín hiệu audio cần được xác định theo SNR dải con của tín hiệu audio này.

Dựa vào cách thức thực hiện khả thi thứ nhất của khía cạnh thứ sáu, theo cách thức thực hiện khả thi thứ hai của khía cạnh thứ sáu, khối xác định

thứ nhất được tạo cấu hình cụ thể để xác định tín hiệu audio này dưới dạng tín hiệu audio cần được xác định trong trường hợp số lượng dải con đầu cao tần, mà nằm trong tín hiệu audio này và có các SNR dải con lớn hơn ngưỡng định trước thứ nhất, là lớn hơn số lượng thứ nhất.

Dựa vào cách thức thực hiện khả thi thứ nhất của khía cạnh thứ sáu, theo cách thức thực hiện khả thi thứ ba của khía cạnh thứ sáu, khối xác định thứ nhất được tạo cấu hình cụ thể để xác định tín hiệu audio này dưới dạng tín hiệu audio cần được xác định trong trường hợp số lượng dải con đầu cao tần, mà nằm trong tín hiệu audio này và có các SNR dải con lớn hơn ngưỡng định trước thứ nhất, là lớn hơn số lượng thứ hai, và số lượng dải con đầu hạ tần, mà nằm trong tín hiệu audio này và có các SNR dải con nhỏ hơn ngưỡng định trước thứ hai, là lớn hơn số lượng thứ ba.

Dựa vào cách thức thực hiện khả thi thứ nhất của khía cạnh thứ sáu, theo cách thức thực hiện khả thi thứ tư của khía cạnh thứ sáu, khối xác định thứ nhất được tạo cấu hình cụ thể để xác định tín hiệu audio này dưới dạng tín hiệu audio cần được xác định trong trường hợp số lượng dải con, mà nằm trong tín hiệu audio này và có giá trị của các SNR dải con lớn hơn ngưỡng định trước thứ ba, là lớn hơn số lượng thứ tư.

Dựa vào khía cạnh thứ sáu, theo cách thức thực hiện khả thi thứ năm của khía cạnh thứ sáu, khối xác định thứ nhất được tạo cấu hình cụ thể để xác định tín hiệu audio này dưới dạng tín hiệu audio cần được xác định trong trường hợp xác định được rằng tín hiệu audio này là tín hiệu âm vô thanh.

Theo phương pháp theo các phương án của sáng chế, nét đặc trưng của tín hiệu audio có thể được xác định, SSNR tăng cường được xác định theo cách tương ứng theo nét đặc trưng của tín hiệu audio, và SSNR tăng cường này được so sánh với ngưỡng quyết định VAD, nên có thể giảm xác suất không phát hiện được tín hiệu có hoạt động.

Mô tả vắn tắt các hình vẽ

Để mô tả các giải pháp kỹ thuật của sáng chế một cách rõ ràng hơn, phần sau đây sẽ mô tả vắn tắt các hình vẽ kèm theo, vốn cần thiết để mô tả các phương án của sáng chế. Các hình vẽ kèm theo trong phần mô tả sau đây chỉ thể hiện một số phương án của sáng chế, và người có kiến thức trung bình trong lĩnh vực này có thể tạo ra các hình vẽ khác dựa vào các hình vẽ kèm theo này mà không cần đến hoạt động có tính sáng tạo nào.

Fig.1 là hình vẽ thể hiện lưu đồ của phương pháp phát hiện tín hiệu audio theo một phương án của sáng chế;

Fig.2 là hình vẽ thể hiện lưu đồ của phương pháp phát hiện tín hiệu audio theo một phương án của sáng chế;

Fig.3 là hình vẽ thể hiện lưu đồ của phương pháp phát hiện tín hiệu audio theo một phương án của sáng chế;

Fig.4 là hình vẽ thể hiện lưu đồ của phương pháp phát hiện tín hiệu audio theo một phương án của sáng chế;

Fig.5 là hình vẽ thể hiện sơ đồ khối cấu trúc của thiết bị theo một phương án của sáng chế;

Fig.6 là hình vẽ thể hiện sơ đồ khối cấu trúc của thiết bị khác theo một phương án của sáng chế;

Fig.7 là hình vẽ thể hiện sơ đồ khối cấu trúc của thiết bị theo một phương án của sáng chế;

Fig.8 là hình vẽ thể hiện sơ đồ khối cấu trúc của thiết bị khác theo một phương án của sáng chế;

Fig.9 là hình vẽ thể hiện sơ đồ khối cấu trúc của thiết bị khác theo một phương án của sáng chế; và

Fig.10 là hình vẽ thể hiện sơ đồ khối cấu trúc của thiết bị khác theo một phương án của sáng chế.

Mô tả chi tiết các phương án thực hiện sáng chế

Phần sau sẽ mô tả rõ các giải pháp kỹ thuật của sáng chế dựa vào các hình vẽ kèm theo và các phương án thực hiện sáng chế. Tất nhiên là các phương án được mô tả chỉ là một số chứ không phải tất cả các phương án của sáng chế. Tất cả các phương án khác mà người có kiến thức trung bình trong lĩnh vực này có thể tạo ra dựa trên các phương án này của sáng chế mà không cần đến hoạt động sáng tạo nào thì cũng nằm trong phạm vi bảo hộ của sáng chế.

Fig.1 là hình vẽ thể hiện lưu đồ của phương pháp phát hiện tín hiệu audio theo một phương án của sáng chế.

101. Xác định tín hiệu audio vào dưới dạng tín hiệu audio cần được xác định.

102. Xác định SSNR tăng cường của tín hiệu audio này, trong đó SSNR tăng cường này lớn hơn SSNR tham chiếu.

103. So sánh SSNR tăng cường này với ngưỡng quyết định VAD để xác định xem tín hiệu audio này có phải là tín hiệu có hoạt động hay không.

Theo phương án này của sáng chế, khi SSNR tăng cường được so sánh với ngưỡng quyết định VAD, thì ngưỡng quyết định VAD tham chiếu có thể được sử dụng, hoặc ngưỡng quyết định VAD đã giảm, mà thu được sau khi ngưỡng quyết định VAD tham chiếu được giảm bằng thuật toán định trước, có thể được sử dụng. Ngưỡng quyết định VAD tham chiếu này có thể là ngưỡng quyết định VAD mặc định, và ngưỡng quyết định VAD tham chiếu này có thể được lưu giữ trước, hoặc có thể là tạm thời thu được qua việc tính toán, trong đó ngưỡng quyết định VAD tham chiếu này có thể được tính toán bằng công nghệ đã biết hiện có. Khi ngưỡng quyết định VAD tham chiếu được giảm bằng thuật toán định trước, thì thuật toán định trước này có thể là nhân ngưỡng quyết định VAD tham chiếu này với hệ số nhỏ hơn 1, hoặc thuật toán khác có thể được sử dụng. Phương án này của sáng

chế không áp đặt giới hạn nào đối với thuật toán cụ thể được sử dụng.

Khi phương pháp tính SSNR thông thường được dùng để tính toán các SSNR của một số tín hiệu audio, thì các SSNR của các tín hiệu audio này có thể thấp hơn ngưỡng quyết định VAD định trước. Tuy nhiên, trên thực tế, các tín hiệu audio này là các tín hiệu audio có hoạt động. Điều này là do những nét đặc trưng của các tín hiệu audio này. Ví dụ, trong trường hợp mà SNR môi trường là tương đối thấp, thì SNR dải con của phần cao tần bị giảm đáng kể. Ngoài ra, do lý thuyết nhận thức âm thanh thường được dùng để thực hiện việc chia dải con, nên SNR dải con của phần cao tần có ảnh hưởng tương đối thấp đến SSNR. Trong trường hợp này, đối với một số tín hiệu, chẳng hạn tín hiệu âm vô thanh, mà có năng lượng chủ yếu tập trung tại phần tần số tương đối cao, thì SSNR thu được sau khi tính toán bằng phương pháp tính SSNR thông thường có thể thấp hơn ngưỡng quyết định VAD, điều này gây ra trường hợp phát hiện trượt tín hiệu có hoạt động. Theo ví dụ khác, đối với một số tín hiệu audio, thì sự phân bố năng lượng của các tín hiệu audio này là tương đối phẳng trên phổ tần nhưng năng lượng tổng thể của các tín hiệu audio này là tương đối thấp. Do đó, trong trường hợp SNR môi trường tương đối thấp, thì SSNR thu được sau khi tính toán bằng phương pháp tính SSNR thông thường có thể thấp hơn ngưỡng quyết định VAD. Theo phương pháp được thể hiện trên Fig.1, cách tăng phù hợp SSNR được sử dụng, để SSNR có thể lớn hơn ngưỡng quyết định VAD. Do đó, xác suất phát hiện trượt tín hiệu có hoạt động có thể được giảm một cách có hiệu quả.

Fig.2 là hình vẽ thể hiện lưu đồ của phương pháp phát hiện tín hiệu audio theo một phương án của sáng chế.

201. Xác định SNR dải con của tín hiệu audio vào.

Phổ tần của tín hiệu audio vào này được chia thành N dải con, trong đó N là số nguyên dương lớn hơn 1. Cụ thể là, lý thuyết nhận thức âm thanh có thể được dùng để chia phổ tần của tín hiệu audio này. Trong trường hợp mà

lý thuyết nhận thức âm thanh được dùng để chia phổ tần của tín hiệu audio, thì độ rộng của dải con càng gần tần số thấp càng hẹp, và độ rộng của dải con càng gần tần số cao càng rộng. Tất nhiên là phổ tần của tín hiệu audio cũng có thể được chia theo cách khác, ví dụ, chia đều phổ tần của tín hiệu audio thành N dải con. SNR dải con của mỗi dải con của tín hiệu audio vào được tính toán, trong đó SNR dải con này là tỉ số giữa năng lượng của dải con trên năng lượng của tạp âm nền ở dải con này. Năng lượng của tạp âm nền trên dải con thường là giá trị được ước lượng thu được sau hoạt động ước lượng của bộ ước lượng tạp âm nền. Cách thức sử dụng bộ ước lượng tạp âm nền để ước lượng năng lượng tạp âm nền tương ứng với mỗi dải con là công nghệ đã được biết rõ trong lĩnh vực này. Do đó, việc này không được mô tả chi tiết ở đây. Người có kiến thức trung bình trong lĩnh vực này có thể thấy rằng SNR dải con có thể là tỉ số năng lượng trực tiếp, hoặc có thể là cách biểu diễn khác của tỉ số năng lượng trực tiếp, chẳng hạn SNR dải con lôgarit. Ngoài ra, người có kiến thức trung bình trong lĩnh vực này có thể còn thấy rằng SNR dải con cũng có thể là SNR dải con thu được sau khi quá trình xử lý tuyến tính hoặc phi tuyến được thực hiện đối với SNR dải con trực tiếp, hoặc có thể là dạng biến đổi khác của SNR dải con. Tỉ số năng lượng trực tiếp của SNR dải con được thể hiện trong công thức sau:

$$snr(k) = E(k) / En(k)$$

Công thức 1.2

trong đó $snr(k)$ biểu thị SNR dải con của dải con thứ k, và $E(k)$ và $En(k)$ lần lượt biểu thị năng lượng của dải con thứ k và năng lượng của tạp âm nền trên dải con thứ k. SNR dải con lôgarit có thể được biểu thị dưới dạng: $snr_{log}(k) = 10 \times \log_{10} snr(k)$, trong đó $snr_{log}(k)$ biểu thị SNR dải con lôgarit của dải con thứ k, và $snr(k)$ biểu thị SNR dải con mà là của dải con thứ k và thu được sau khi tính toán bằng công thức 1.2. Người có kiến thức trung bình trong lĩnh vực này có thể còn thấy rằng năng lượng dải con được dùng để tính toán SNR dải con có thể là năng lượng của tín hiệu audio vào trên

dải con, hoặc có thể là năng lượng thu được sau khi năng lượng của tạp âm nền trên dải con được trừ đi từ năng lượng của tín hiệu audio vào trên dải con. Việc tính toán SNR là phù hợp mà không nằm ngoài ý nghĩa của SNR.

202. Xác định tín hiệu audio vào dưới dạng tín hiệu audio cần được xác định.

Tùy ý, theo một phương án, bước xác định tín hiệu audio vào dưới dạng tín hiệu audio cần được xác định có thể là bước: xác định tín hiệu audio này dưới dạng tín hiệu audio cần được xác định theo SNR dải con mà là của tín hiệu audio này và xác định được ở bước 201.

Tùy ý, theo một phương án, trong trường hợp tín hiệu audio này được xác định dưới dạng tín hiệu audio cần được xác định theo SNR dải con của tín hiệu audio này, thì bước xác định tín hiệu audio vào dưới dạng tín hiệu audio cần được xác định là bước: xác định tín hiệu audio này dưới dạng tín hiệu audio cần được xác định trong trường hợp số lượng dải con đầu cao tần, mà nằm trong tín hiệu audio này và có các SNR dải con lớn hơn ngưỡng định trước thứ nhất, là lớn hơn số lượng thứ nhất.

Tùy ý, theo phương án khác, trong trường hợp tín hiệu audio này được xác định dưới dạng tín hiệu audio cần được xác định theo SNR dải con của tín hiệu audio này, thì bước xác định tín hiệu audio vào dưới dạng tín hiệu audio cần được xác định là bước: xác định tín hiệu audio này dưới dạng tín hiệu audio cần được xác định trong trường hợp số lượng dải con đầu cao tần, mà nằm trong tín hiệu audio này và có các SNR dải con lớn hơn ngưỡng định trước thứ nhất, là lớn hơn số lượng thứ hai, và số lượng dải con đầu hạ tần, mà nằm trong tín hiệu audio này và có các SNR dải con nhỏ hơn ngưỡng định trước thứ hai, là lớn hơn số lượng thứ ba. Theo phương án này của sáng chế, đầu cao tần và đầu hạ tần của một khung của tín hiệu audio là tương đối, tức là, phần có tần số tương đối cao là đầu cao tần, và phần có tần số tương đối thấp là đầu hạ tần.

Tùy ý, theo phương án khác, trong trường hợp tín hiệu audio này được

xác định dưới dạng tín hiệu audio cần được xác định theo SNR dải con của tín hiệu audio này, thì bước xác định tín hiệu audio vào dưới dạng tín hiệu audio cần được xác định là bước: xác định tín hiệu audio này dưới dạng tín hiệu audio cần được xác định trong trường hợp số lượng dải con, mà nằm trong tín hiệu audio này và có giá trị của các SNR dải con lớn hơn ngưỡng định trước thứ ba, là lớn hơn số lượng thứ tư.

Ngưỡng định trước thứ nhất và ngưỡng định trước thứ hai có thể thu được bằng phương pháp tập hợp thống kê theo một lượng lớn mẫu tiếng nói. Cụ thể là, số liệu thống kê về các SNR dải con của các dải con đầu cao tần được tập hợp trong một lượng lớn các mẫu tiếng nói âm vô thanh bao gồm tạp âm nền, và ngưỡng định trước thứ nhất được xác định theo các SNR dải con này, sao cho các SNR dải con của hầu hết trong số các dải con đầu cao tần trong các mẫu âm vô thanh này là lớn hơn ngưỡng định trước thứ nhất này. Tương tự, số liệu thống kê về các SNR dải con của các dải con đầu hạ tần là được tập hợp trong các mẫu tiếng nói âm vô thanh này, và ngưỡng định trước thứ hai được xác định theo các SNR dải con này, sao cho các SNR dải con của hầu hết trong số các dải con đầu hạ tần trong các mẫu tiếng nói âm vô thanh này là nhỏ hơn ngưỡng định trước thứ hai này.

Ngưỡng định trước thứ ba cũng được thu thập bằng phương pháp tập hợp thống kê. Cụ thể là, ngưỡng định trước thứ ba được xác định theo các SNR dải con của một lượng lớn tín hiệu tạp âm, sao cho các SNR dải con của hầu hết trong số các dải con trong các tín hiệu tạp âm này là nhỏ hơn ngưỡng định trước thứ ba này.

Số lượng thứ nhất, số lượng thứ hai, số lượng thứ ba, và số lượng thứ tư cũng được thu thập bằng phương pháp tập hợp thống kê. Số lượng thứ nhất được lấy làm ví dụ, trong đó trong một lượng lớn khung mẫu tiếng nói âm vô thanh bao gồm tạp âm, thì số liệu thống kê về lượng dải con của các dải con đầu cao tần, mà có các SNR dải con lớn hơn ngưỡng định trước thứ nhất, được tập hợp, và số lượng thứ nhất này được xác định theo lượng này,

sao cho số lượng dải con đầu cao tần, mà nằm trong hầu hết trong số các khung mẫu tiếng nói âm vô thanh này và có các SNR dải con lớn hơn ngưỡng định trước thứ nhất, là lớn hơn số lượng thứ nhất. Phương pháp thu thập số lượng thứ hai là tương tự như phương pháp thu thập số lượng thứ nhất. Số lượng thứ hai có thể giống với số lượng thứ nhất, hoặc số lượng thứ hai có thể khác với số lượng thứ nhất. Tương tự, đối với số lượng thứ ba, thì trong một lượng lớn khung mẫu tiếng nói âm vô thanh bao gồm tạp âm, số liệu thống kê về lượng dải con của các dải con đầu hạ tần, mà có các SNR dải con nhỏ hơn ngưỡng định trước thứ hai, được tập hợp, và số lượng thứ ba này được xác định theo lượng này, sao cho số lượng dải con đầu hạ tần, mà nằm trong hầu hết trong số các khung mẫu tiếng nói âm vô thanh này và có các SNR dải con nhỏ hơn ngưỡng định trước thứ hai, là lớn hơn số lượng thứ ba này. Đối với số lượng thứ tư, trong một lượng lớn khung tín hiệu tạp âm, số liệu thống kê về số lượng dải con, mà có các SNR dải con nhỏ hơn ngưỡng định trước thứ ba, được tập hợp, và số lượng thứ tư này được xác định theo lượng này, sao cho số lượng dải con, mà nằm trong hầu hết trong số các khung mẫu tạp âm này và có các SNR dải con nhỏ hơn ngưỡng định trước thứ ba, là lớn hơn số lượng thứ tư này.

Tùy ý, theo phương án khác, việc tín hiệu audio vào có phải là tín hiệu audio cần được xác định hay không có thể được xác định bằng cách xác định xem tín hiệu audio vào có phải là tín hiệu âm vô thanh hay không. Trong trường hợp này, SNR dải con của tín hiệu audio này không cần phải được xác định khi việc tín hiệu audio này có phải là tín hiệu audio cần được xác định hay không đang được xác định. Nói cách khác, bước 201 không cần phải được thực hiện khi việc tín hiệu audio này có phải là tín hiệu audio cần được xác định hay không đang được xác định. Cụ thể là, bước xác định tín hiệu audio vào dưới dạng tín hiệu audio cần được xác định là bước: xác định tín hiệu audio này dưới dạng tín hiệu audio cần được xác định trong trường hợp xác định được rằng tín hiệu audio này là tín hiệu âm vô thanh.

Cụ thể là, người có kiến thức trung bình trong lĩnh vực này có thể thấy rằng có thể có nhiều phương pháp để phát hiện xem tín hiệu audio này có phải là tín hiệu âm vô thanh hay không. Ví dụ, việc tín hiệu audio này có phải là tín hiệu âm vô thanh hay không có thể được xác định bằng cách phát hiện tỉ lệ qua điểm không (Zero-Crossing Rate - ZCR) trong miền thời gian của tín hiệu audio này. Cụ thể là, trong trường hợp mà ZCR của tín hiệu audio này lớn hơn ngưỡng ZCR, thì tín hiệu audio này được xác định là tín hiệu âm vô thanh, trong đó ngưỡng ZCR này được xác định theo một lượng lớn thực nghiệm.

203. Xác định SSNR tăng cường của tín hiệu audio này, trong đó SSNR tăng cường này lớn hơn SSNR tham chiếu.

SSNR tham chiếu có thể là SSNR thu được sau khi tính toán bằng công thức 1.1. Có thể thấy từ công thức 1.1 rằng quy trình lấy trọng số không được thực hiện đối với SNR dải con của dải con bất kì khi SSNR tham chiếu đang được tính, tức là, các trọng số của các SNR dải con của tất cả các dải con là bằng nhau khi SSNR tham chiếu đang được tính.

Tuỳ ý, theo một phương án, trong trường hợp mà số lượng dải con đầu cao tần, mà nằm trong tín hiệu audio và có các SNR dải con lớn hơn, ngưỡng định trước thứ nhất, là lớn hơn số lượng thứ nhất, hoặc trong trường hợp mà số lượng dải con đầu cao tần, mà nằm trong tín hiệu audio này và có các SNR dải con lớn hơn ngưỡng định trước thứ nhất, là lớn hơn số lượng thứ hai, và số lượng dải con đầu hạ tần, mà nằm trong tín hiệu audio này và có các SNR dải con nhỏ hơn ngưỡng định trước thứ hai, là lớn hơn số lượng thứ ba, thì bước xác định SSNR tăng cường của tín hiệu audio này là các bước: xác định trọng số của SNR dải con của mỗi dải con trong tín hiệu audio này, trong đó trọng số của dải con đầu cao tần mà có SNR dải con lớn hơn ngưỡng định trước thứ nhất là lớn hơn trọng số của SNR dải con của dải con khác; và xác định SSNR tăng cường theo SNR dải con của mỗi dải con và trọng số của SNR dải con của mỗi dải con trong tín hiệu audio này.

Ví dụ, nếu tín hiệu audio này được chia thành 20 dải con, tức là dải con 0 đến dải con 19, theo lý thuyết nhận thức âm thanh, và các tỉ số tín hiệu trên tạp âm của dải con 18 và dải con 19 đều lớn hơn trị số định trước thứ nhất T1, thì bốn dải con, tức là dải con 20 đến dải con 23, có thể được thêm vào. Cụ thể là, dải con 18 và dải con 19, mà có các tỉ số tín hiệu trên tạp âm lớn hơn T1, có thể lần lượt được chia thành dải con 18a, dải con 18b, và dải con 18c; và dải con 19a, dải con 19b, và dải con 19c. Trong trường hợp này, dải con 18 có thể được coi là dải con mẹ của dải con 18a, dải con 18b, và dải con 18c, và dải con 19 có thể được coi là dải con mẹ của dải con 19a, dải con 19b, và dải con 19c. Các trị số của các tỉ số tín hiệu trên tạp âm của dải con 18a, dải con 18b, và dải con 18c là giống như trị số của tỉ số tín hiệu trên tạp âm của dải con mẹ của chúng, và các trị số của các tỉ số tín hiệu trên tạp âm của dải con 19a, dải con 19b, và dải con 19c là giống như trị số của tỉ số tín hiệu trên tạp âm của dải con mẹ của chúng. Theo cách này, 20 dải con, mà thu được ban đầu sau khi chia, được chia lại thành 24 dải con. Bởi vì VAD vẫn được thiết kế theo 20 dải con trong quá trình phát hiện tín hiệu có hoạt động, nên 24 dải con này cần được ánh xạ trở lại 20 dải con này để xác định SSNR tăng cường. Tóm lại, khi SSNR tăng cường được xác định bằng cách tăng số lượng dải con đầu cao tần mà có các SNR dải con lớn hơn ngưỡng định trước thứ nhất, thì việc tính toán có thể được thực hiện bằng công thức sau:

$$SSNR' = \frac{20}{24} \times \left[2 \times (snr(18) + snr(19)) + \sum_{k=0}^{19} snr(k) \right] \quad \text{Công thức 1.3}$$

trong đó $SSNR'$ biểu thị SSNR tăng cường, và $snr(k)$ biểu thị SNR dải con của dải con thứ k.

Nếu SSNR thu được sau khi tính toán bằng công thức 1.1 là SSNR tham chiếu, thì SSNR tham chiếu thu được sau khi tính toán là $\sum_{k=0}^{19} snr(k)$. Rõ ràng là đối với tín hiệu audio thuộc loại thứ nhất, thì giá trị của SSNR

tăng cường thu được sau khi tính toán bằng công thức 1.3 là lớn hơn giá trị của SSNR tham chiếu thu được sau khi tính toán bằng công thức 1.1.

Theo ví dụ khác, nếu tín hiệu audio này được chia thành 20 dải con, tức là dải con 0 đến dải con 19, theo lý thuyết nhận thức âm thanh, $snr(18)$ và $snr(19)$ đều lớn hơn trị số định trước thứ nhất T1, và $snr(0)$ đến $snr(17)$ đều nhỏ hơn ngưỡng định trước thứ hai T2, thì SSNR tăng cường có thể được xác định bằng công thức sau:

$$SSNR' = a_1 \times snr(18) + a_2 \times snr(19) + \sum_{k=0}^{17} snr(k) \quad \text{Công thức 1.4}$$

trong đó $SSNR'$ biểu thị SSNR tăng cường, $snr(k)$ biểu thị SNR dải con của dải con thứ k, a_1 và a_2 là các thông số tăng trọng số, và các giá trị của a_1 và a_2 làm cho $a_1 \times snr(18) + a_2 \times snr(19)$ lớn hơn $snr(18) + snr(19)$. Rõ ràng là giá trị của SSNR tăng cường thu được sau khi tính toán bằng công thức 1.4 là lớn hơn giá trị của SSNR tham chiếu thu được sau khi tính toán bằng công thức 1.1.

Tuỳ ý, theo phương án khác, bước xác định SSNR tăng cường của tín hiệu audio là bước: xác định SSNR tham chiếu của tín hiệu audio này, và xác định SSNR tăng cường theo SSNR tham chiếu của tín hiệu audio này.

Tuỳ ý, SSNR tăng cường có thể được xác định bằng công thức sau:

$$SSNR' = x * SSNR + y \quad \text{Công thức 1.5}$$

trong đó SSNR là SSNR tham chiếu của tín hiệu audio, $SSNR'$ biểu thị SSNR tăng cường, và x và y biểu thị các thông số tăng cường. Ví dụ, trị số của x có thể là 1,05, và trị số của y có thể là 1. Người có kiến thức trung bình trong lĩnh vực này có thể thấy rằng các trị số của x và y có thể là các giá trị phù hợp khác mà làm cho SSNR tăng cường lớn hơn SSNR tham chiếu một cách phù hợp.

Tuỳ ý, SSNR tăng cường có thể được xác định bằng công thức sau:

$$SSNR' = f(x) * SSNR + h(y)$$

Công thức 1.6

trong đó SSNR biểu thị SSNR ban đầu của tín hiệu audio, $SSNR'$ biểu thị SSNR tăng cường, và $f(x)$ và $h(y)$ biểu thị các hàm tăng cường. Ví dụ, $f(x)$ và $h(y)$ có thể là các hàm liên quan đến tỉ số tín hiệu trên tạp âm dài hạn (Long-term Signal-to-Noise Ratio - LSNR) của tín hiệu audio, trong đó LSNR của tín hiệu audio là SNR trung bình hoặc SNR có trọng số trong một khoảng thời gian tương đối dài. Ví dụ, khi $lsnr$ lớn hơn 20, thì $f(lsnr)$ có thể bằng 1,1, và $y(lsnr)$ có thể bằng 2; khi $lsnr$ nhỏ hơn 20 và lớn hơn 15, thì $f(lsnr)$ có thể bằng 1,05, và $y(lsnr)$ có thể bằng 1; và khi $lsnr$ nhỏ hơn 15, thì $f(lsnr)$ có thể bằng 1, và $y(lsnr)$ có thể bằng 0. Người có kiến thức trung bình trong lĩnh vực này có thể thấy rằng $f(x)$ và $h(y)$ có thể có các dạng phù hợp khác mà làm cho SSNR tăng cường lớn hơn SSNR tham chiếu một cách phù hợp.

204. So sánh SSNR tăng cường này với ngưỡng quyết định VAD để xác định xem tín hiệu audio này có phải là tín hiệu có hoạt động hay không.

Cụ thể là, khi SSNR tăng cường được so sánh với ngưỡng quyết định VAD, nếu SSNR tăng cường này lớn hơn ngưỡng quyết định VAD, thì tín hiệu audio này được xác định là tín hiệu có hoạt động; hoặc nếu SSNR tăng cường này không lớn hơn ngưỡng quyết định VAD, thì tín hiệu audio này được xác định là tín hiệu không có hoạt động.

Tuỳ ý, theo phương án khác, trước bước so sánh SSNR tăng cường với ngưỡng quyết định VAD, thì phương pháp này có thể còn bao gồm bước: dùng thuật toán định trước để giảm ngưỡng quyết định VAD, để thu được ngưỡng quyết định VAD đã giảm. Trong trường hợp này, bước so sánh SSNR tăng cường với ngưỡng quyết định VAD cụ thể là bước: so sánh SSNR tăng cường với ngưỡng quyết định VAD đã giảm để xác định xem tín hiệu audio này có phải là tín hiệu có hoạt động hay không. Ngưỡng quyết định VAD tham chiếu có thể là ngưỡng quyết định VAD mặc định, và

ngưỡng quyết định VAD tham chiếu này có thể được lưu giữ trước, hoặc có thể là tạm thời thu được qua việc tính toán, trong đó ngưỡng quyết định VAD tham chiếu này có thể được tính toán bằng công nghệ đã biết hiện có. Khi ngưỡng quyết định VAD tham chiếu được giảm bằng thuật toán định trước, thì thuật toán định trước này có thể là nhân ngưỡng quyết định VAD tham chiếu này với hệ số nhỏ hơn 1, hoặc thuật toán khác có thể được sử dụng. Phương án này của sáng chế không áp đặt giới hạn nào đối với thuật toán cụ thể được sử dụng. Ngưỡng quyết định VAD có thể được giảm một cách phù hợp bằng thuật toán định trước, sao cho SSNR tăng cường trở nên lớn hơn ngưỡng quyết định VAD đã giảm. Do đó, xác suất phát hiện trượt tín hiệu có hoạt động có thể được giảm.

Theo phương pháp được thể hiện trên Fig.2, nét đặc trưng của tín hiệu audio được xác định, SSNR tăng cường được xác định theo cách tương ứng theo nét đặc trưng này của tín hiệu audio, và SSNR tăng cường được so sánh với ngưỡng quyết định VAD. Theo cách này, xác suất phát hiện trượt tín hiệu có hoạt động có thể được giảm.

Fig.3 là hình vẽ thể hiện lưu đồ của phương pháp phát hiện tín hiệu audio theo một phương án của sáng chế.

301. Xác định tín hiệu audio vào dưới dạng tín hiệu audio cần được xác định.

302. Xác định trọng số của SNR dải con của mỗi dải con trong tín hiệu audio này, trong đó trọng số của SNR dải con của dải con đầu cao tần mà có SNR dải con lớn hơn ngưỡng định trước thứ nhất là lớn hơn trọng số của SNR dải con của dải con khác.

303. Xác định SSNR tăng cường theo SNR dải con của mỗi dải con và trọng số của SNR dải con của mỗi dải con trong tín hiệu audio này, trong đó SSNR tăng cường này là lớn hơn SSNR tham chiếu.

SSNR tham chiếu có thể là SSNR thu được sau khi tính toán bằng công thức 1.1. Có thể thấy từ công thức 1.1 rằng quy trình lấy trọng số không

được thực hiện đối với SNR dải con của dải con bất kì khi SSNR tham chiếu đang được tính, tức là, các trọng số của các SNR dải con của tất cả các dải con là bằng nhau khi SSNR tham chiếu đang được tính.

Ví dụ, nếu tín hiệu audio này được chia thành 20 dải con, tức là dải con 0 đến dải con 19, theo lý thuyết nhận thức âm thanh, và các tỉ số tín hiệu trên tạp âm của dải con 18 và dải con 19 đều lớn hơn trị số định trước thứ nhất T1, thì bốn dải con, tức là dải con 20 đến dải con 23, có thể được thêm vào. Cụ thể là, dải con 18 và dải con 19, mà có các tỉ số tín hiệu trên tạp âm lớn hơn T1, có thể lần lượt được chia thành dải con 18a, dải con 18b, và dải con 18c; và dải con 19a, dải con 19b, và dải con 19c. Trong trường hợp này, dải con 18 có thể được coi là dải con mẹ của dải con 18a, dải con 18b, và dải con 18c, và dải con 19 có thể được coi là dải con mẹ của dải con 19a, dải con 19b, và dải con 19c. Các trị số của các tỉ số tín hiệu trên tạp âm của dải con 18a, dải con 18b, và dải con 18c là giống như trị số của tỉ số tín hiệu trên tạp âm của dải con mẹ của chúng, và các trị số của các tỉ số tín hiệu trên tạp âm của dải con 19a, dải con 19b, và dải con 19c là giống như trị số của tỉ số tín hiệu trên tạp âm của dải con mẹ của chúng. Theo cách này, 20 dải con, mà thu được ban đầu sau khi chia, được chia lại thành 24 dải con. Bởi vì VAD vẫn được thiết kế theo 20 dải con trong quá trình phát hiện tín hiệu có hoạt động, nên 24 dải con này cần được ánh xạ trở lại 20 dải con này để xác định SSNR tăng cường. Tóm lại, khi SSNR tăng cường được xác định bằng cách tăng số lượng dải con đầu cao tần mà có các SNR dải con lớn hơn ngưỡng định trước thứ nhất, thì việc tính toán có thể được thực hiện bằng công thức sau:

$$SSNR' = \frac{20}{24} \times \left[2 \times (snr(18) + snr(19)) + \sum_{k=0}^{19} snr(k) \right] \quad \text{Công thức 1.3}$$

trong đó $SSNR'$ biểu thị SSNR tăng cường, và $snr(k)$ biểu thị SNR dải con của dải con thứ k.

Nếu SSNR thu được sau khi tính toán bằng công thức 1.1 là SSNR

tham chiếu, thì SSNR tham chiếu thu được sau khi tính toán là $\sum_{k=0}^{19} snr(k)$. Rõ ràng là đối với tín hiệu audio thuộc loại thứ nhất, thì giá trị của SSNR tăng cường thu được sau khi tính toán bằng công thức 1.3 là lớn hơn giá trị của SSNR tham chiếu thu được sau khi tính toán bằng công thức 1.1.

Theo ví dụ khác, nếu tín hiệu audio này được chia thành 20 dải con, tức là dải con 0 đến dải con 19, theo lý thuyết nhận thức âm thanh, $snr(18)$ và $snr(19)$ đều lớn hơn trị số định trước thứ nhất T1, và $snr(0)$ đến $snr(17)$ đều nhỏ hơn ngưỡng định trước thứ hai T2, thì SSNR tăng cường có thể được xác định bằng công thức sau:

$$SSNR' = a_1 \times snr(18) + a_2 \times snr(19) + \sum_{k=0}^{17} snr(k) \quad \text{Công thức 1.4}$$

trong đó $SSNR'$ biểu thị SSNR tăng cường, $snr(k)$ biểu thị SNR dải con của dải con thứ k, a_1 và a_2 là các thông số tăng trọng số, và các giá trị của a_1 và a_2 làm cho $a_1 \times snr(18) + a_2 \times snr(19)$ lớn hơn $snr(18) + snr(19)$. Rõ ràng là giá trị của SSNR tăng cường thu được sau khi tính toán bằng công thức 1.4 là lớn hơn giá trị của SSNR tham chiếu thu được sau khi tính toán bằng công thức 1.1.

304. So sánh SSNR tăng cường này với ngưỡng quyết định VAD để xác định xem tín hiệu audio này có phải là tín hiệu có hoạt động hay không.

Cụ thể là, khi SSNR tăng cường được so sánh với ngưỡng quyết định VAD, nếu SSNR tăng cường này lớn hơn ngưỡng quyết định VAD, thì tín hiệu audio này được xác định là tín hiệu có hoạt động; hoặc nếu SSNR tăng cường này không lớn hơn ngưỡng quyết định VAD, thì tín hiệu audio này được xác định là tín hiệu không có hoạt động.

Theo phương pháp được thể hiện trên Fig.3, nét đặc trưng của tín hiệu audio có thể được xác định, SSNR tăng cường được xác định theo cách tương ứng theo nét đặc trưng này của tín hiệu audio, và SSNR tăng cường

được so sánh với ngưỡng quyết định VAD. Do đó, xác suất phát hiện trượt tín hiệu có hoạt động có thể được giảm.

Ngoài ra, bước xác định tín hiệu audio vào dưới dạng tín hiệu audio cần được xác định là bước: xác định tín hiệu audio này dưới dạng tín hiệu audio cần được xác định theo SNR dải con của tín hiệu audio này.

Tùy ý, theo một phương án, trong trường hợp tín hiệu audio này được xác định dưới dạng tín hiệu audio cần được xác định theo SNR dải con của tín hiệu audio này, thì bước xác định tín hiệu audio này dưới dạng tín hiệu audio cần được xác định là bước: xác định tín hiệu audio này dưới dạng tín hiệu audio cần được xác định trong trường hợp số lượng dải con đầu cao tần, mà nằm trong tín hiệu audio này và có các SNR dải con lớn hơn ngưỡng định trước thứ nhất, là lớn hơn số lượng thứ nhất.

Tùy ý, theo phương án khác, trong trường hợp tín hiệu audio này được xác định dưới dạng tín hiệu audio cần được xác định theo SNR dải con của tín hiệu audio này, thì bước xác định tín hiệu audio này dưới dạng tín hiệu audio cần được xác định là bước: xác định tín hiệu audio này dưới dạng tín hiệu audio cần được xác định trong trường hợp số lượng dải con đầu cao tần, mà nằm trong tín hiệu audio này và có các SNR dải con lớn hơn ngưỡng định trước thứ nhất, là lớn hơn số lượng thứ hai, và số lượng dải con đầu hạ tần, mà nằm trong tín hiệu audio này và có các SNR dải con nhỏ hơn ngưỡng định trước thứ hai, là lớn hơn số lượng thứ ba.

Ngưỡng định trước thứ nhất và ngưỡng định trước thứ hai có thể thu được bằng phương pháp tập hợp thống kê theo một lượng lớn mẫu tiếng nói. Cụ thể là, số liệu thống kê về các SNR dải con của các dải con đầu cao tần được tập hợp trong một lượng lớn các mẫu tiếng nói âm vô thanh bao gồm tập âm nền, và ngưỡng định trước thứ nhất được xác định theo các SNR dải con này, sao cho các SNR dải con của hầu hết trong số các dải con đầu cao tần trong các mẫu âm vô thanh này là lớn hơn ngưỡng định trước thứ nhất này. Tương tự, số liệu thống kê về các SNR dải con của các dải con đầu hạ

tần là được tập hợp trong các mẫu tiếng nói âm vô thanh này, và ngưỡng định trước thứ hai được xác định theo các SNR dải con này, sao cho các SNR dải con của hầu hết trong số các dải con đầu hạ tần trong các mẫu tiếng nói âm vô thanh này là nhỏ hơn ngưỡng định trước thứ hai này.

Số lượng thứ nhất, số lượng thứ hai, và số lượng thứ ba cũng được thu thập bằng phương pháp tập hợp thống kê. Số lượng thứ nhất được lấy làm ví dụ, trong đó trong một lượng lớn khung mẫu tiếng nói âm vô thanh bao gồm tạp âm, thì số liệu thống kê về lượng dải con của các dải con đầu cao tần, mà có các SNR dải con lớn hơn ngưỡng định trước thứ nhất, được tập hợp, và số lượng thứ nhất này được xác định theo lượng này, sao cho số lượng dải con đầu cao tần, mà nằm trong hầu hết trong số các khung mẫu tiếng nói âm vô thanh này và có các SNR dải con lớn hơn ngưỡng định trước thứ nhất, là lớn hơn số lượng thứ nhất. Phương pháp thu thập số lượng thứ hai là tương tự như phương pháp thu thập số lượng thứ nhất. Số lượng thứ hai có thể giống với số lượng thứ nhất, hoặc số lượng thứ hai có thể khác với số lượng thứ nhất. Tương tự, đối với số lượng thứ ba, thì trong một lượng lớn khung mẫu tiếng nói âm vô thanh bao gồm tạp âm, số liệu thống kê về lượng dải con của các dải con đầu hạ tần, mà có các SNR dải con nhỏ hơn ngưỡng định trước thứ hai, được tập hợp, và số lượng thứ ba này được xác định theo lượng này, sao cho số lượng dải con đầu hạ tần, mà nằm trong hầu hết trong số các khung mẫu tiếng nói âm vô thanh này và có các SNR dải con nhỏ hơn ngưỡng định trước thứ hai, là lớn hơn số lượng thứ ba này.

Theo các phương án được thể hiện trên các hình vẽ từ Fig.1 đến Fig.3, việc tín hiệu audio vào có phải là tín hiệu có hoạt động hay không là được xác định theo cách sử dụng SSNR tăng cường. Theo phương pháp được thể hiện trên Fig.4, việc tín hiệu audio vào có phải là tín hiệu có hoạt động hay không được xác định theo cách là giảm ngưỡng quyết định VAD.

Fig.4 là hình vẽ thể hiện lưu đồ của phương pháp phát hiện tín hiệu audio theo một phương án của sáng chế.

401. Xác định tín hiệu audio vào dưới dạng tín hiệu audio cần được xác định.

Tùy ý, theo một phương án, bước xác định tín hiệu audio vào dưới dạng tín hiệu audio cần được xác định là bước: xác định tín hiệu audio này dưới dạng tín hiệu audio cần được xác định theo SNR dải con mà là của tín hiệu audio này và xác định được ở bước 201.

Tùy ý, theo một phương án, trong trường hợp tín hiệu audio này được xác định dưới dạng tín hiệu audio cần được xác định theo SNR dải con của tín hiệu audio này, thì bước xác định tín hiệu audio vào dưới dạng tín hiệu audio cần được xác định là bước: xác định tín hiệu audio này dưới dạng tín hiệu audio cần được xác định trong trường hợp số lượng dải con đầu cao tần, mà nằm trong tín hiệu audio này và có các SNR dải con lớn hơn ngưỡng định trước thứ nhất, là lớn hơn số lượng thứ nhất.

Tùy ý, theo phương án khác, trong trường hợp tín hiệu audio này được xác định dưới dạng tín hiệu audio cần được xác định theo SNR dải con của tín hiệu audio này, thì bước xác định tín hiệu audio vào dưới dạng tín hiệu audio cần được xác định là bước: xác định tín hiệu audio này dưới dạng tín hiệu audio cần được xác định trong trường hợp số lượng dải con đầu cao tần, mà nằm trong tín hiệu audio này và có các SNR dải con lớn hơn ngưỡng định trước thứ nhất, là lớn hơn số lượng thứ hai, và số lượng dải con đầu hạ tần, mà nằm trong tín hiệu audio này và có các SNR dải con nhỏ hơn ngưỡng định trước thứ hai, là lớn hơn số lượng thứ ba.

Tùy ý, theo phương án khác, trong trường hợp tín hiệu audio này được xác định dưới dạng tín hiệu audio cần được xác định theo SNR dải con của tín hiệu audio này, thì bước xác định tín hiệu audio vào dưới dạng tín hiệu audio cần được xác định là bước: xác định tín hiệu audio này dưới dạng tín hiệu audio cần được xác định trong trường hợp số lượng dải con, mà nằm trong tín hiệu audio này và có giá trị của các SNR dải con lớn hơn ngưỡng định trước thứ ba, là lớn hơn số lượng thứ tư.

Ngưỡng định trước thứ nhất và ngưỡng định trước thứ hai có thể thu được bằng phương pháp tập hợp thống kê theo một lượng lớn mẫu tiếng nói. Cụ thể là, số liệu thống kê về các SNR dải con của các dải con đầu cao tần được tập hợp trong một lượng lớn các mẫu tiếng nói âm vô thanh bao gồm tạp âm nền, và ngưỡng định trước thứ nhất được xác định theo các SNR dải con này, sao cho các SNR dải con của hầu hết trong số các dải con đầu cao tần trong các mẫu âm vô thanh này là lớn hơn ngưỡng định trước thứ nhất này. Tương tự, số liệu thống kê về các SNR dải con của các dải con đầu hạ tần là được tập hợp trong các mẫu tiếng nói âm vô thanh này, và ngưỡng định trước thứ hai được xác định theo các SNR dải con này, sao cho các SNR dải con của hầu hết trong số các dải con đầu hạ tần trong các mẫu tiếng nói âm vô thanh này là nhỏ hơn ngưỡng định trước thứ hai này.

Ngưỡng định trước thứ ba cũng được thu thập bằng phương pháp tập hợp thống kê. Cụ thể là, ngưỡng định trước thứ ba được xác định theo các SNR dải con của một lượng lớn tín hiệu tạp âm, sao cho các SNR dải con của hầu hết trong số các dải con trong các tín hiệu tạp âm này là nhỏ hơn ngưỡng định trước thứ ba này.

Số lượng thứ nhất, số lượng thứ hai, số lượng thứ ba, và số lượng thứ tư cũng được thu thập bằng phương pháp tập hợp thống kê. Số lượng thứ nhất được lấy làm ví dụ, trong đó trong một lượng lớn khung mẫu tiếng nói âm vô thanh bao gồm tạp âm, thì số liệu thống kê về lượng dải con của các dải con đầu cao tần, mà có các SNR dải con lớn hơn ngưỡng định trước thứ nhất, được tập hợp, và số lượng thứ nhất này được xác định theo lượng này, sao cho số lượng dải con đầu cao tần, mà nằm trong hầu hết trong số các khung mẫu tiếng nói âm vô thanh này và có các SNR dải con lớn hơn ngưỡng định trước thứ nhất, là lớn hơn số lượng thứ nhất. Phương pháp thu thập số lượng thứ hai là tương tự như phương pháp thu thập số lượng thứ nhất. Số lượng thứ hai có thể giống với số lượng thứ nhất, hoặc số lượng thứ hai có thể khác với số lượng thứ nhất. Tương tự, đối với số lượng thứ ba, thì

trong một lượng lớn khung mẫu tiếng nói âm vô thanh bao gồm tạp âm, số liệu thống kê về lượng dải con của các dải con đầu hạ tần, mà có các SNR dải con nhỏ hơn ngưỡng định trước thứ hai, được tập hợp, và số lượng thứ ba này được xác định theo lượng này, sao cho số lượng dải con đầu hạ tần, mà nằm trong hầu hết trong số các khung mẫu tiếng nói âm vô thanh này và có các SNR dải con nhỏ hơn ngưỡng định trước thứ hai, là lớn hơn số lượng thứ ba này. Đối với số lượng thứ tư, trong một lượng lớn khung tín hiệu tạp âm, số liệu thống kê về số lượng dải con, mà có các SNR dải con nhỏ hơn ngưỡng định trước thứ ba, được tập hợp, và số lượng thứ tư này được xác định theo lượng này, sao cho số lượng dải con, mà nằm trong hầu hết trong số các khung mẫu tạp âm này và có các SNR dải con nhỏ hơn ngưỡng định trước thứ ba, là lớn hơn số lượng thứ tư này.

Tuỳ ý, theo phương án khác, việc tín hiệu audio vào có phải là tín hiệu audio cần được xác định hay không có thể được xác định bằng cách xác định xem tín hiệu audio vào có phải là tín hiệu âm vô thanh hay không. Trong trường hợp này, SNR dải con của tín hiệu audio này không cần phải được xác định khi việc tín hiệu audio này có phải là tín hiệu audio cần được xác định hay không đang được xác định. Nói cách khác, bước 201 không cần phải được thực hiện khi việc tín hiệu audio này có phải là tín hiệu audio cần được xác định hay không đang được xác định. Cụ thể là, bước xác định tín hiệu audio vào dưới dạng tín hiệu audio cần được xác định là bước: xác định tín hiệu audio này dưới dạng tín hiệu audio cần được xác định trong trường hợp xác định được rằng tín hiệu audio này là tín hiệu âm vô thanh. Cụ thể là, người có kiến thức trung bình trong lĩnh vực này có thể thấy rằng có thể có nhiều phương pháp để phát hiện xem tín hiệu audio này có phải là tín hiệu âm vô thanh hay không. Ví dụ, việc tín hiệu audio này có phải là tín hiệu âm vô thanh hay không có thể được xác định bằng cách phát hiện ZCR trong miền thời gian của tín hiệu audio này. Cụ thể là, trong trường hợp mà ZCR của tín hiệu audio này lớn hơn ngưỡng ZCR, thì tín hiệu audio này

được xác định là tín hiệu âm vô thanh, trong đó ngưỡng ZCR này được xác định theo một lượng lớn thực nghiệm.

402. Thu thập SSNR tham chiếu của tín hiệu audio này.

Cụ thể là, SSNR tham chiếu có thể là SSNR thu được sau khi tính toán bằng công thức 1.1.

403. Dùng thuật toán định trước để giảm ngưỡng quyết định VAD tham chiếu, để thu được ngưỡng quyết định VAD đã giảm.

Cụ thể là, ngưỡng quyết định VAD tham chiếu này có thể là ngưỡng quyết định VAD mặc định, và ngưỡng quyết định VAD tham chiếu này có thể được lưu giữ trước, hoặc có thể là tạm thời thu được qua việc tính toán, trong đó ngưỡng quyết định VAD tham chiếu này có thể được tính toán bằng công nghệ đã biết hiện có. Khi ngưỡng quyết định VAD tham chiếu được giảm bằng thuật toán định trước, thì thuật toán định trước này có thể là nhân ngưỡng quyết định VAD tham chiếu này với hệ số nhỏ hơn 1, hoặc thuật toán khác có thể được sử dụng. Phương án này của sáng chế không áp đặt giới hạn nào đối với thuật toán cụ thể được sử dụng. Ngưỡng quyết định VAD có thể được giảm một cách phù hợp bằng thuật toán định trước, sao cho SSNR tăng cường trở nên lớn hơn ngưỡng quyết định VAD đã giảm. Do đó, xác suất phát hiện trượt tín hiệu có hoạt động có thể được giảm.

404. So sánh SSNR tham chiếu với ngưỡng quyết định VAD đã giảm để xác định xem tín hiệu audio này có phải là tín hiệu có hoạt động hay không.

Khi phương pháp tính SSNR thông thường được dùng để tính toán các SSNR của một số tín hiệu audio, thì các SSNR của các tín hiệu audio này có thể thấp hơn ngưỡng quyết định VAD định trước. Tuy nhiên, trên thực tế, các tín hiệu audio này là các tín hiệu audio có hoạt động. Điều này là do những nét đặc trưng của các tín hiệu audio này. Ví dụ, trong trường hợp mà SNR môi trường là tương đối thấp, thì SNR dải con của phần cao tần bị giảm đáng kể. Ngoài ra, do lý thuyết nhận thức âm thanh thường được dùng

để thực hiện việc chia dải con, nên SNR dải con của phần cao tần có ảnh hưởng tương đối thấp đến SSNR. Trong trường hợp này, đối với một số tín hiệu, chẳng hạn tín hiệu âm vô thanh, mà có năng lượng chủ yếu tập trung tại phần tần số tương đối cao, thì SSNR thu được sau khi tính toán bằng phương pháp tính SSNR thông thường có thể thấp hơn ngưỡng quyết định VAD, điều này gây ra trường hợp phát hiện trượt tín hiệu có hoạt động. Theo ví dụ khác, đối với một số tín hiệu audio, thì sự phân bố năng lượng của các tín hiệu audio này là tương đối phẳng trên phổ tần nhưng năng lượng tổng thể của các tín hiệu audio này là tương đối thấp. Do đó, trong trường hợp SNR môi trường tương đối thấp, thì SSNR thu được sau khi tính toán bằng phương pháp tính SSNR thông thường có thể thấp hơn ngưỡng quyết định VAD. Theo phương pháp trên Fig.4, cách giảm ngưỡng quyết định VAD được sử dụng, nên SSNR thu được sau khi tính toán bằng phương pháp tính SSNR thông thường là lớn hơn ngưỡng quyết định VAD. Do đó, xác suất phát hiện trượt tín hiệu có hoạt động có thể được giảm một cách có hiệu quả.

Fig.5 là hình vẽ thể hiện sơ đồ khối cấu trúc của thiết bị theo một phương án của sáng chế. Thiết bị trên Fig.5 có thể thực hiện tất cả các bước được thể hiện trên Fig.1 hoặc Fig.2. Như được thể hiện trên Fig.5, thiết bị 500 bao gồm khối xác định thứ nhất 501, khối xác định thứ hai 502, và khối xác định thứ ba 503.

Khối xác định thứ nhất 501 được tạo cấu hình để xác định tín hiệu audio vào dưới dạng tín hiệu audio cần được xác định.

Khối xác định thứ hai 502 được tạo cấu hình để xác định tỉ số tín hiệu trên tạp âm theo phân đoạn (Segmental Signal-to-Noise Ratio - SSNR) tăng cường của tín hiệu audio, trong đó SSNR tăng cường này lớn hơn SSNR tham chiếu.

Khối xác định thứ ba 503 được tạo cấu hình để so sánh SSNR tăng cường này với ngưỡng quyết định phát hiện hoạt động thoại (Voice Activity

Detection - VAD) để xác định xem tín hiệu audio này có phải là tín hiệu có hoạt động hay không.

Thiết bị 500 trên Fig.5 có thể xác định nét đặc trưng của tín hiệu audio vào, xác định SSNR tăng cường theo cách tương ứng theo nét đặc trưng này của tín hiệu audio, và so sánh SSNR tăng cường này với ngưỡng quyết định VAD, nên tỉ lệ phát hiện trượt tín hiệu có hoạt động có thể được giảm.

Tùy ý, theo một phương án, khối xác định thứ nhất 501 được tạo cấu hình cụ thể để xác định tín hiệu audio này dưới dạng tín hiệu audio cần được xác định theo SNR dải con của tín hiệu audio này.

Tùy ý, theo một phương án, trong trường hợp khối xác định thứ nhất 501 xác định tín hiệu audio này dưới dạng tín hiệu audio cần được xác định theo SNR dải con của tín hiệu audio này, thì khối xác định thứ nhất 501 được tạo cấu hình cụ thể để xác định tín hiệu audio này dưới dạng tín hiệu audio cần được xác định trong trường hợp số lượng dải con đầu cao tần, mà nằm trong tín hiệu audio này và có các SNR dải con lớn hơn ngưỡng định trước thứ nhất, là lớn hơn số lượng thứ nhất.

Tùy ý, theo phương án khác, trong trường hợp khối xác định thứ nhất 501 xác định tín hiệu audio này dưới dạng tín hiệu audio cần được xác định theo SNR dải con của tín hiệu audio này, thì khối xác định thứ nhất 501 được tạo cấu hình cụ thể để xác định tín hiệu audio này dưới dạng tín hiệu audio cần được xác định trong trường hợp số lượng dải con đầu cao tần, mà nằm trong tín hiệu audio này và có các SNR dải con lớn hơn ngưỡng định trước thứ nhất, là lớn hơn số lượng thứ hai, và số lượng dải con đầu hạ tần, mà nằm trong tín hiệu audio này và có các SNR dải con nhỏ hơn ngưỡng định trước thứ hai, là lớn hơn số lượng thứ ba.

Tùy ý, theo phương án khác, trong trường hợp khối xác định thứ nhất 501 xác định tín hiệu audio này dưới dạng tín hiệu audio cần được xác định theo SNR dải con của tín hiệu audio này, thì khối xác định thứ nhất 501 được tạo cấu hình cụ thể để xác định tín hiệu audio này dưới dạng tín hiệu

audio cần được xác định trong trường hợp số lượng dải con, mà nằm trong tín hiệu audio này và có giá trị của các SNR dải con lớn hơn ngưỡng định trước thứ ba, là lớn hơn số lượng thứ tư.

Tuỳ ý, theo phương án khác, khối xác định thứ nhất 501 được tạo cấu hình cụ thể để xác định tín hiệu audio này dưới dạng tín hiệu audio cần được xác định trong trường hợp xác định được rằng tín hiệu audio này là tín hiệu âm vô thanh. Cụ thể là, người có kiến thức trung bình trong lĩnh vực này có thể thấy rằng có thể có nhiều phương pháp để phát hiện xem tín hiệu audio này có phải là tín hiệu âm vô thanh hay không. Ví dụ, việc tín hiệu audio này có phải là tín hiệu âm vô thanh hay không có thể được xác định bằng cách phát hiện ZCR trong miền thời gian của tín hiệu audio này. Cụ thể là, trong trường hợp mà ZCR của tín hiệu audio này lớn hơn ngưỡng ZCR, thì tín hiệu audio này được xác định là tín hiệu âm vô thanh, trong đó ngưỡng ZCR này được xác định theo một lượng lớn thực nghiệm.

Ngưỡng định trước thứ nhất và ngưỡng định trước thứ hai có thể thu được bằng phương pháp tập hợp thống kê theo một lượng lớn mẫu tiếng nói. Cụ thể là, số liệu thống kê về các SNR dải con của các dải con đầu cao tần được tập hợp trong một lượng lớn các mẫu tiếng nói âm vô thanh bao gồm tạp âm nền, và ngưỡng định trước thứ nhất được xác định theo các SNR dải con này, sao cho các SNR dải con của hầu hết trong số các dải con đầu cao tần trong các mẫu âm vô thanh này là lớn hơn ngưỡng định trước thứ nhất này. Tương tự, số liệu thống kê về các SNR dải con của các dải con đầu hạ tần là được tập hợp trong các mẫu tiếng nói âm vô thanh này, và ngưỡng định trước thứ hai được xác định theo các SNR dải con này, sao cho các SNR dải con của hầu hết trong số các dải con đầu hạ tần trong các mẫu tiếng nói âm vô thanh này là nhỏ hơn ngưỡng định trước thứ hai này.

Ngưỡng định trước thứ ba cũng được thu thập bằng phương pháp tập hợp thống kê. Cụ thể là, ngưỡng định trước thứ ba được xác định theo các SNR dải con của một lượng lớn tín hiệu tạp âm, sao cho các SNR dải con

của hầu hết trong số các dải con trong các tín hiệu tạp âm này là nhỏ hơn ngưỡng định trước thứ ba này.

Số lượng thứ nhất, số lượng thứ hai, số lượng thứ ba, và số lượng thứ tư cũng được thu thập bằng phương pháp tập hợp thống kê. Số lượng thứ nhất được lấy làm ví dụ, trong đó trong một lượng lớn mẫu tiếng nói bao gồm tạp âm, thì số liệu thống kê về lượng dải con của các dải con đầu cao tần, mà có các SNR dải con lớn hơn ngưỡng định trước thứ nhất, được tập hợp, và số lượng thứ nhất này được xác định theo lượng này, sao cho số lượng dải con đầu cao tần, mà nằm trong hầu hết trong số các mẫu tiếng nói này và có các SNR dải con lớn hơn ngưỡng định trước thứ nhất, là lớn hơn số lượng thứ nhất. Phương pháp để xác định số lượng thứ hai là tương tự như phương pháp để xác định số lượng thứ nhất. Số lượng thứ hai có thể giống với số lượng thứ nhất, hoặc có thể khác với số lượng thứ nhất. Tương tự, đối với số lượng thứ ba, thì trong một lượng lớn mẫu tiếng nói bao gồm tạp âm, số liệu thống kê về lượng dải con của các dải con đầu hạ tần, mà có các SNR dải con lớn hơn ngưỡng định trước thứ hai, được tập hợp, và số lượng thứ ba này được xác định theo lượng này, sao cho số lượng dải con đầu hạ tần, mà nằm trong hầu hết trong số các mẫu tiếng nói này và có các SNR dải con lớn hơn ngưỡng định trước thứ hai, là lớn hơn số lượng thứ ba này. Đối với số lượng thứ tư, trong lượng lớn mẫu tiếng nói bao gồm tạp âm này, thì số liệu thống kê về số lượng dải con, mà có các SNR dải con lớn hơn ngưỡng định trước thứ ba, được tập hợp, và số lượng thứ tư này được xác định theo lượng này, sao cho số lượng dải con, mà nằm trong hầu hết trong số các mẫu tiếng nói này và có các SNR dải con lớn hơn ngưỡng định trước thứ ba, là lớn hơn số lượng thứ tư này.

Ngoài ra, khối xác định thứ hai 502 được tạo cấu hình cụ thể để xác định trọng số của SNR dải con của mỗi dải con trong tín hiệu audio, trong đó trọng số của dải con đầu cao tần, mà có SNR dải con lớn hơn ngưỡng định trước thứ nhất, là lớn hơn trọng số của SNR dải con của dải con khác,

và xác định SSNR tăng cường theo SNR của mỗi dải con và trọng số của SNR dải con của mỗi dải con trong tín hiệu audio này.

Tuỳ ý, theo một phương án, khối xác định thứ hai 502 được tạo cấu hình cụ thể để xác định SSNR tham chiếu của tín hiệu audio, và xác định SSNR tăng cường theo SSNR tham chiếu của tín hiệu audio này.

SSNR tham chiếu có thể là SSNR thu được sau khi tính toán bằng công thức 1.1. Khi SSNR tham chiếu đang được tính, thì các trọng số của các SNR dải con mà là của tất cả các dải con và được bao gồm trong SSNR này là giống nhau trong SSNR này.

Tuỳ ý, theo phương án khác, khối xác định thứ hai 502 được tạo cấu hình cụ thể để xác định SSNR tăng cường bằng công thức sau:

$$SSNR' = x * SSNR + y \quad \text{Công thức 1.7}$$

trong đó SSNR biểu thị SSNR tham chiếu, SSNR' biểu thị SSNR tăng cường, và x và y biểu thị các thông số tăng cường. Ví dụ, trị số của x có thể là 1,05, và trị số của y có thể là 1. Người có kiến thức trung bình trong lĩnh vực này có thể thấy rằng các trị số của x và y có thể là các giá trị phù hợp khác mà làm cho SSNR tăng cường lớn hơn SSNR tham chiếu một cách phù hợp.

Tuỳ ý, theo phương án khác, khối xác định thứ hai 502 được tạo cấu hình cụ thể để xác định SSNR tăng cường bằng công thức sau:

$$SSNR' = f(x) * SSNR + h(y) \quad \text{Công thức 1.8}$$

trong đó SSNR biểu thị SSNR tham chiếu, SSNR' biểu thị SSNR tăng cường, và f(x) và h(y) biểu thị các hàm tăng cường. Ví dụ, f(x) và h(y) có thể là các hàm liên quan đến LSNR của tín hiệu audio, trong đó LSNR của tín hiệu audio là SNR trung bình hoặc SNR có trọng số trong một khoảng thời gian tương đối dài. Ví dụ, khi lsnr lớn hơn 20, thì f(lsnr) có thể bằng 1,1, và y(lsnr) có thể bằng 2; khi lsnr nhỏ hơn 20 và lớn hơn 15, thì f(lsnr) có thể bằng 1,05, và y(lsnr) có thể bằng 1; và khi lsnr nhỏ hơn 15, thì f(lsnr) có thể

bằng 1, và $y(\text{lsnr})$ có thể bằng 0. Người có kiến thức trung bình trong lĩnh vực này có thể thấy rằng $f(x)$ và $h(y)$ có thể có các dạng phù hợp khác mà làm cho SSNR tăng cường lớn hơn SSNR tham chiếu một cách phù hợp.

Khối xác định thứ ba 503 được tạo cấu hình cụ thể để so sánh SSNR tăng cường với ngưỡng quyết định VAD để xác định, theo kết quả so sánh, xem tín hiệu audio này có phải là tín hiệu có hoạt động hay không. Cụ thể là, nếu SSNR tăng cường này lớn hơn ngưỡng quyết định VAD, thì tín hiệu audio này được xác định là tín hiệu có hoạt động, hoặc nếu SSNR tăng cường này nhỏ hơn ngưỡng quyết định VAD, thì tín hiệu audio này được xác định là tín hiệu không có hoạt động.

Tùy ý, theo phương án khác, thuật toán định trước cũng có thể được dùng để giảm ngưỡng quyết định VAD tham chiếu để thu được ngưỡng quyết định VAD đã giảm, và ngưỡng quyết định VAD đã giảm này được dùng để xác định xem tín hiệu audio này có phải là tín hiệu có hoạt động hay không. Trong trường hợp này, thiết bị 500 có thể còn bao gồm khối xác định thứ tư 504, trong đó khối xác định thứ tư 504 này được tạo cấu hình để dùng thuật toán định trước để giảm ngưỡng quyết định VAD, để thu được ngưỡng quyết định VAD đã giảm. Trong trường hợp này, khối xác định thứ ba 503 được tạo cấu hình cụ thể để so sánh SSNR tăng cường với ngưỡng quyết định VAD đã giảm để xác định xem tín hiệu audio này có phải là tín hiệu có hoạt động hay không.

Fig.6 là hình vẽ thể hiện sơ đồ khối cấu trúc của thiết bị khác theo một phương án của sáng chế. Thiết bị trên Fig.6 có thể thực hiện tất cả các bước được thể hiện trên Fig.3. Như được thể hiện trên Fig.6, thiết bị 600 bao gồm khối xác định thứ nhất 601, khối xác định thứ hai 602, và khối xác định thứ ba 603.

Khối xác định thứ nhất 601 được tạo cấu hình để xác định tín hiệu audio vào dưới dạng tín hiệu audio cần được xác định.

Khối xác định thứ hai 602 được tạo cấu hình để xác định trọng số của

SNR dải con của mỗi dải con trong tín hiệu audio này, trong đó trọng số của SNR dải con của dải con đầu cao tần, mà có SNR dải con lớn hơn ngưỡng định trước thứ nhất, là lớn hơn trọng số của SNR dải con của dải con khác, và xác định SSNR tăng cường theo SNR dải con của mỗi dải con và trọng số của SNR dải con của mỗi dải con trong tín hiệu audio này, trong đó SSNR tăng cường này là lớn hơn SSNR tham chiếu.

Khối xác định thứ ba 603 được tạo cấu hình để so sánh SSNR tăng cường này với ngưỡng quyết định VAD để xác định xem tín hiệu audio này có phải là tín hiệu có hoạt động hay không.

Thiết bị 600 trên Fig.6 có thể xác định nét đặc trưng của tín hiệu audio vào, xác định SSNR tăng cường theo cách tương ứng theo nét đặc trưng này của tín hiệu audio, và so sánh SSNR tăng cường này với ngưỡng quyết định VAD, nên tỉ lệ phát hiện trượt tín hiệu có hoạt động có thể được giảm.

Ngoài ra, khối xác định thứ nhất 601 được tạo cấu hình cụ thể để xác định tín hiệu audio này dưới dạng tín hiệu audio cần được xác định theo SNR dải con của tín hiệu audio này.

Tuỳ ý, theo một phương án, khối xác định thứ nhất 601 được tạo cấu hình cụ thể để xác định tín hiệu audio này dưới dạng tín hiệu audio cần được xác định trong trường hợp số lượng dải con đầu cao tần, mà nằm trong tín hiệu audio này và có các SNR dải con lớn hơn ngưỡng định trước thứ nhất, là lớn hơn số lượng thứ nhất.

Tuỳ ý, theo phương án khác, khối xác định thứ nhất 601 được tạo cấu hình cụ thể để xác định tín hiệu audio này dưới dạng tín hiệu audio cần được xác định trong trường hợp số lượng dải con đầu cao tần, mà nằm trong tín hiệu audio này và có các SNR dải con lớn hơn ngưỡng định trước thứ nhất, là lớn hơn số lượng thứ hai, và số lượng dải con đầu hạ tần, mà nằm trong tín hiệu audio này và có các SNR dải con nhỏ hơn ngưỡng định trước thứ hai, là lớn hơn số lượng thứ ba.

Ngưỡng định trước thứ nhất và ngưỡng định trước thứ hai có thể thu

được bằng phương pháp tập hợp thống kê theo một lượng lớn mẫu tiếng nói. Cụ thể là, số liệu thống kê về các SNR dải con của các dải con đầu cao tần được tập hợp trong một lượng lớn các mẫu tiếng nói âm vô thanh bao gồm tạp âm nền, và ngưỡng định trước thứ nhất được xác định theo các SNR dải con này, sao cho các SNR dải con của hầu hết trong số các dải con đầu cao tần trong các mẫu âm vô thanh này là lớn hơn ngưỡng định trước thứ nhất này. Tương tự, số liệu thống kê về các SNR dải con của các dải con đầu hạ tần là được tập hợp trong các mẫu tiếng nói âm vô thanh này, và ngưỡng định trước thứ hai được xác định theo các SNR dải con này, sao cho các SNR dải con của hầu hết trong số các dải con đầu hạ tần trong các mẫu tiếng nói âm vô thanh này là nhỏ hơn ngưỡng định trước thứ hai này.

Số lượng thứ nhất, số lượng thứ hai, và số lượng thứ ba cũng được thu thập bằng phương pháp tập hợp thống kê. Số lượng thứ nhất được lấy làm ví dụ, trong đó trong một lượng lớn khung mẫu tiếng nói âm vô thanh bao gồm tạp âm, thì số liệu thống kê về lượng dải con của các dải con đầu cao tần, mà có các SNR dải con lớn hơn ngưỡng định trước thứ nhất, được tập hợp, và số lượng thứ nhất này được xác định theo lượng này, sao cho số lượng dải con đầu cao tần, mà nằm trong hầu hết trong số các khung mẫu tiếng nói âm vô thanh này và có các SNR dải con lớn hơn ngưỡng định trước thứ nhất, là lớn hơn số lượng thứ nhất. Phương pháp thu thập số lượng thứ hai là tương tự như phương pháp thu thập số lượng thứ nhất. Số lượng thứ hai có thể giống với số lượng thứ nhất, hoặc số lượng thứ hai có thể khác với số lượng thứ nhất. Tương tự, đối với số lượng thứ ba, thì trong một lượng lớn khung mẫu tiếng nói âm vô thanh bao gồm tạp âm, số liệu thống kê về lượng dải con của các dải con đầu hạ tần, mà có các SNR dải con nhỏ hơn ngưỡng định trước thứ hai, được tập hợp, và số lượng thứ ba này được xác định theo lượng này, sao cho số lượng dải con đầu hạ tần, mà nằm trong hầu hết trong số các khung mẫu tiếng nói âm vô thanh này và có các SNR dải con nhỏ hơn ngưỡng định trước thứ hai, là lớn hơn số lượng thứ ba này.

Fig.7 là hình vẽ thể hiện sơ đồ khối cấu trúc của thiết bị theo một phương án của sáng chế. Thiết bị trên Fig.7 có thể thực hiện tất cả các bước được thể hiện trên Fig.1 hoặc Fig.2. Như được thể hiện trên Fig.7, thiết bị 700 bao gồm bộ xử lý 701 và bộ nhớ 702. Bộ xử lý 701 có thể là bộ xử lý thông dụng, bộ xử lý tín hiệu số (Digital Signal Processor - DSP), mạch tích hợp chuyên dụng (Application Specific Integrated Circuit - ASIC), mảng cổng lập trình được dạng trường (Field Programmable Gate Array - FPGA) hoặc thành phần logic lập trình được khác, cổng rời rạc hoặc thành phần logic tranzito, hoặc thành phần phần cứng rời rạc, mà có thể thực hiện các phương pháp, các bước, và các sơ đồ khối logic đã được bộc lộ theo các phương án của sáng chế. Bộ xử lý thông dụng có thể là bộ vi xử lý, hoặc bộ xử lý nêu trên có thể là bộ xử lý thông thường bất kì hoặc các phương tiện tương tự. Các bước của các phương pháp đã được bộc lộ theo các phương án của sáng chế có thể được thực hiện trực tiếp bởi bộ xử lý giải mã phần cứng, hoặc được thực hiện bởi tổ hợp của các môđun phần cứng và phần mềm trong bộ xử lý giải mã. Môđun phần mềm này có thể được đặt trong phương tiện lưu trữ đã biết trong lĩnh vực, chẳng hạn bộ nhớ truy cập ngẫu nhiên (Random Access Memory - RAM), bộ nhớ flash, bộ nhớ chỉ đọc (Read-Only Memory - ROM), bộ nhớ chỉ đọc lập trình được, bộ nhớ lập trình và xoá được bằng điện, hoặc thanh ghi. Phương tiện lưu trữ này được đặt trong bộ nhớ 702. Bộ xử lý 701 đọc ra lệnh từ bộ nhớ 702, và thực hiện các bước của các phương pháp nêu trên cùng với phần cứng.

Bộ xử lý 701 được tạo cấu hình để xác định tín hiệu audio vào dưới dạng tín hiệu audio cần được xác định.

Bộ xử lý 701 được tạo cấu hình để xác định SSNR tăng cường của tín hiệu audio này, trong đó SSNR tăng cường này lớn hơn SSNR tham chiếu.

Bộ xử lý 701 được tạo cấu hình để so sánh SSNR tăng cường này với ngưỡng quyết định VAD để xác định xem tín hiệu audio này có phải là tín hiệu có hoạt động hay không.

Thiết bị 700 trên Fig.7 có thể xác định nét đặc trưng của tín hiệu audio vào, xác định SSNR tăng cường theo cách tương ứng theo nét đặc trưng này của tín hiệu audio, và so sánh SSNR tăng cường này với ngưỡng quyết định VAD, nên tỉ lệ phát hiện trượt tín hiệu có hoạt động có thể được giảm.

Tuỳ ý, theo một phương án, bộ xử lý 701 được tạo cấu hình cụ thể để xác định tín hiệu audio này dưới dạng tín hiệu audio cần được xác định theo SNR dải con của tín hiệu audio này.

Tuỳ ý, theo một phương án, trong trường hợp bộ xử lý 701 xác định tín hiệu audio này dưới dạng tín hiệu audio cần được xác định theo SNR dải con của tín hiệu audio này, thì bộ xử lý 701 được tạo cấu hình cụ thể để xác định tín hiệu audio này dưới dạng tín hiệu audio cần được xác định trong trường hợp số lượng dải con đầu cao tần, mà nằm trong tín hiệu audio này và có các SNR dải con lớn hơn ngưỡng định trước thứ nhất, là lớn hơn số lượng thứ nhất.

Tuỳ ý, theo phương án khác, trong trường hợp bộ xử lý 701 xác định tín hiệu audio này dưới dạng tín hiệu audio cần được xác định theo SNR dải con của tín hiệu audio này, thì bộ xử lý 701 được tạo cấu hình cụ thể để xác định tín hiệu audio này dưới dạng tín hiệu audio cần được xác định trong trường hợp số lượng dải con đầu cao tần, mà nằm trong tín hiệu audio này và có các SNR dải con lớn hơn ngưỡng định trước thứ nhất, là lớn hơn số lượng thứ hai, và số lượng dải con đầu hạ tần, mà nằm trong tín hiệu audio này và có các SNR dải con nhỏ hơn ngưỡng định trước thứ hai, là lớn hơn số lượng thứ ba.

Tuỳ ý, theo phương án khác, trong trường hợp bộ xử lý 701 xác định tín hiệu audio này dưới dạng tín hiệu audio cần được xác định theo SNR dải con của tín hiệu audio này, thì bộ xử lý 701 được tạo cấu hình cụ thể để xác định tín hiệu audio này dưới dạng tín hiệu audio cần được xác định trong trường hợp số lượng dải con, mà nằm trong tín hiệu audio này và có giá trị của các SNR dải con lớn hơn ngưỡng định trước thứ ba, là lớn hơn số lượng

thứ tư.

Tuỳ ý, theo phương án khác, bộ xử lý 701 được tạo cấu hình cụ thể để xác định tín hiệu audio này dưới dạng tín hiệu audio cần được xác định trong trường hợp xác định được rằng tín hiệu audio này là tín hiệu âm vô thanh. Cụ thể là, người có kiến thức trung bình trong lĩnh vực này có thể thấy rằng có thể có nhiều phương pháp để phát hiện xem tín hiệu audio này có phải là tín hiệu âm vô thanh hay không. Ví dụ, việc tín hiệu audio này có phải là tín hiệu âm vô thanh hay không có thể được xác định bằng cách phát hiện ZCR trong miền thời gian của tín hiệu audio này. Cụ thể là, trong trường hợp mà ZCR của tín hiệu audio này lớn hơn ngưỡng ZCR, thì tín hiệu audio này được xác định là tín hiệu âm vô thanh, trong đó ngưỡng ZCR này được xác định theo một lượng lớn thực nghiệm.

Ngưỡng định trước thứ nhất và ngưỡng định trước thứ hai có thể thu được bằng phương pháp tập hợp thống kê theo một lượng lớn mẫu tiếng nói. Cụ thể là, số liệu thống kê về các SNR dải con của các dải con đầu cao tần được tập hợp trong một lượng lớn các mẫu tiếng nói âm vô thanh bao gồm tạp âm nền, và ngưỡng định trước thứ nhất được xác định theo các SNR dải con này, sao cho các SNR dải con của hầu hết trong số các dải con đầu cao tần trong các mẫu âm vô thanh này là lớn hơn ngưỡng định trước thứ nhất này. Tương tự, số liệu thống kê về các SNR dải con của các dải con đầu hạ tần là được tập hợp trong các mẫu tiếng nói âm vô thanh này, và ngưỡng định trước thứ hai được xác định theo các SNR dải con này, sao cho các SNR dải con của hầu hết trong số các dải con đầu hạ tần trong các mẫu tiếng nói âm vô thanh này là nhỏ hơn ngưỡng định trước thứ hai này.

Ngưỡng định trước thứ ba cũng được thu thập bằng phương pháp tập hợp thống kê. Cụ thể là, ngưỡng định trước thứ ba được xác định theo các SNR dải con của một lượng lớn tín hiệu tạp âm, sao cho các SNR dải con của hầu hết trong số các dải con trong các tín hiệu tạp âm này là nhỏ hơn ngưỡng định trước thứ ba này.

Số lượng thứ nhất, số lượng thứ hai, số lượng thứ ba, và số lượng thứ tư cũng được thu thập bằng phương pháp tập hợp thống kê. Số lượng thứ nhất được lấy làm ví dụ, trong đó trong một lượng lớn mẫu tiếng nói bao gồm tạp âm, thì số liệu thống kê về lượng dải con của các dải con đầu cao tần, mà có các SNR dải con lớn hơn ngưỡng định trước thứ nhất, được tập hợp, và số lượng thứ nhất này được xác định theo lượng này, sao cho số lượng dải con đầu cao tần, mà nằm trong hầu hết trong số các mẫu tiếng nói này và có các SNR dải con lớn hơn ngưỡng định trước thứ nhất, là lớn hơn số lượng thứ nhất. Phương pháp để xác định số lượng thứ hai là tương tự như phương pháp để xác định số lượng thứ nhất. Số lượng thứ hai có thể giống với số lượng thứ nhất, hoặc có thể khác với số lượng thứ nhất. Tương tự, đối với số lượng thứ ba, thì trong một lượng lớn mẫu tiếng nói bao gồm tạp âm, số liệu thống kê về lượng dải con của các dải con đầu hạ tần, mà có các SNR dải con lớn hơn ngưỡng định trước thứ hai, được tập hợp, và số lượng thứ ba này được xác định theo lượng này, sao cho số lượng dải con đầu hạ tần, mà nằm trong hầu hết trong số các mẫu tiếng nói này và có các SNR dải con lớn hơn ngưỡng định trước thứ hai, là lớn hơn số lượng thứ ba này. Đối với số lượng thứ tư, trong lượng lớn mẫu tiếng nói bao gồm tạp âm này, thì số liệu thống kê về số lượng dải con, mà có các SNR dải con lớn hơn ngưỡng định trước thứ ba, được tập hợp, và số lượng thứ tư này được xác định theo lượng này, sao cho số lượng dải con, mà nằm trong hầu hết trong số các mẫu tiếng nói này và có các SNR dải con lớn hơn ngưỡng định trước thứ ba, là lớn hơn số lượng thứ tư này.

Ngoài ra, bộ xử lý 701 được tạo cấu hình cụ thể để xác định trọng số của SNR dải con của mỗi dải con trong tín hiệu audio, trong đó trọng số của dải con đầu cao tần, mà có SNR dải con lớn hơn ngưỡng định trước thứ nhất, là lớn hơn trọng số của SNR dải con của dải con khác, và xác định SSNR tăng cường theo SNR của mỗi dải con và trọng số của SNR dải con của mỗi dải con trong tín hiệu audio này.

Tuỳ ý, theo một phương án, bộ xử lý 701 được tạo cấu hình cụ thể để xác định SSNR tham chiếu của tín hiệu audio, và xác định SSNR tăng cường theo SSNR tham chiếu của tín hiệu audio này.

SSNR tham chiếu có thể là SSNR thu được sau khi tính toán bằng công thức 1.1. Khi SSNR tham chiếu đang được tính, thì các trọng số của các SNR dải con mà là của tất cả các dải con và được bao gồm trong SSNR này là giống nhau trong SSNR này.

Tuỳ ý, theo phương án khác, bộ xử lý 701 được tạo cấu hình cụ thể để xác định SSNR tăng cường bằng công thức sau:

$$SSNR' = x * SSNR + y \quad \text{Công thức 1.7}$$

trong đó SSNR biểu thị SSNR tham chiếu, $SSNR'$ biểu thị SSNR tăng cường, và x và y biểu thị các thông số tăng cường. Ví dụ, trị số của x có thể là 1,07, và trị số của y có thể là 1. Người có kiến thức trung bình trong lĩnh vực này có thể thấy rằng các trị số của x và y có thể là các giá trị phù hợp khác mà làm cho SSNR tăng cường lớn hơn SSNR tham chiếu một cách phù hợp.

Tuỳ ý, theo phương án khác, bộ xử lý 701 được tạo cấu hình cụ thể để xác định SSNR tăng cường bằng công thức sau:

$$SSNR' = f(x) * SSNR + h(y) \quad \text{Công thức 1.8}$$

trong đó SSNR biểu thị SSNR tham chiếu, $SSNR'$ biểu thị SSNR tăng cường, và $f(x)$ và $h(y)$ biểu thị các hàm tăng cường. Ví dụ, $f(x)$ và $h(y)$ có thể là các hàm liên quan đến LSNR của tín hiệu audio, trong đó LSNR của tín hiệu audio là SNR trung bình hoặc SNR có trọng số trong một khoảng thời gian tương đối dài. Ví dụ, khi $lsnr$ lớn hơn 20, thì $f(lsnr)$ có thể bằng 1,1, và $y(lsnr)$ có thể bằng 2; khi $lsnr$ nhỏ hơn 20 và lớn hơn 17, thì $f(lsnr)$ có thể bằng 1,07, và $y(lsnr)$ có thể bằng 1; và khi $lsnr$ nhỏ hơn 17, thì $f(lsnr)$ có thể bằng 1, và $y(lsnr)$ có thể bằng 0. Người có kiến thức trung bình trong lĩnh vực này có thể thấy rằng $f(x)$ và $h(y)$ có thể có các dạng phù hợp khác mà

làm cho SSNR tăng cường lớn hơn SSNR tham chiếu một cách phù hợp.

Bộ xử lý 701 được tạo cấu hình cụ thể để so sánh SSNR tăng cường với ngưỡng quyết định VAD để xác định, theo kết quả so sánh, xem tín hiệu audio này có phải là tín hiệu có hoạt động hay không. Cụ thể là, nếu SSNR tăng cường này lớn hơn ngưỡng quyết định VAD, thì tín hiệu audio này được xác định là tín hiệu có hoạt động, hoặc nếu SSNR tăng cường này nhỏ hơn ngưỡng quyết định VAD, thì tín hiệu audio này được xác định là tín hiệu không có hoạt động.

Tùy ý, theo phương án khác, thuật toán định trước cũng có thể được dùng để giảm ngưỡng quyết định VAD tham chiếu để thu được ngưỡng quyết định VAD đã giảm, và ngưỡng quyết định VAD đã giảm này được dùng để xác định xem tín hiệu audio này có phải là tín hiệu có hoạt động hay không. Trong trường hợp này, bộ xử lý 701 có thể còn được tạo cấu hình để dùng thuật toán định trước để giảm ngưỡng quyết định VAD, để thu được ngưỡng quyết định VAD đã giảm. Trong trường hợp này, bộ xử lý 701 được tạo cấu hình cụ thể để so sánh SSNR tăng cường với ngưỡng quyết định VAD đã giảm để xác định xem tín hiệu audio này có phải là tín hiệu có hoạt động hay không.

Fig.8 là hình vẽ thể hiện sơ đồ khối cấu trúc của thiết bị khác theo một phương án của sáng chế. Thiết bị trên Fig.8 có thể thực hiện tất cả các bước được thể hiện trên Fig.3. Như được thể hiện trên Fig.8, thiết bị 800 bao gồm bộ xử lý 801 và bộ nhớ 802. Bộ xử lý 801 có thể là bộ xử lý thông dụng, DSP, ASIC, FPGA hoặc thành phần logic lập trình được khác, cổng rời rạc hoặc thành phần logic tranzito, hoặc thành phần phần cứng rời rạc, mà có thể thực hiện các phương pháp, các bước, và các sơ đồ khối logic đã được bộc lộ theo các phương án của sáng chế. Bộ xử lý thông dụng có thể là bộ vi xử lý, hoặc bộ xử lý nêu trên có thể là bộ xử lý thông thường bất kì hoặc, v.v.. Các bước của các phương pháp đã được bộc lộ theo các phương án của sáng chế có thể được thực hiện trực tiếp bởi bộ xử lý giải mã phần cứng,

hoặc được thực hiện bởi tổ hợp của các môđun phần cứng và phần mềm trong bộ xử lý giải mã. Môđun phần mềm này có thể được đặt trong phương tiện lưu trữ đã biết trong lĩnh vực, chẳng hạn RAM, bộ nhớ flash, ROM, bộ nhớ chỉ đọc lập trình được, bộ nhớ lập trình và xoá được bằng điện, hoặc thanh ghi. Phương tiện lưu trữ này được đặt trong bộ nhớ 802. Bộ xử lý 801 đọc ra lệnh từ bộ nhớ 802, và thực hiện các bước của các phương pháp nêu trên cùng với phần cứng.

Bộ xử lý 801 được tạo cấu hình để xác định tín hiệu audio vào dưới dạng tín hiệu audio cần được xác định.

Bộ xử lý 801 được tạo cấu hình để xác định trọng số của SNR dải con của mỗi dải con trong tín hiệu audio này, trong đó trọng số của SNR dải con của dải con đầu cao tần, mà có SNR dải con lớn hơn ngưỡng định trước thứ nhất, là lớn hơn trọng số của SNR dải con của dải con khác, và xác định SSNR tăng cường theo SNR dải con của mỗi dải con và trọng số của SNR dải con của mỗi dải con trong tín hiệu audio này, trong đó SSNR tăng cường này là lớn hơn SSNR tham chiếu.

Bộ xử lý 801 được tạo cấu hình để so sánh SSNR tăng cường này với ngưỡng quyết định VAD để xác định xem tín hiệu audio này có phải là tín hiệu có hoạt động hay không.

Thiết bị 800 trên Fig.8 có thể xác định nét đặc trưng của tín hiệu audio vào, xác định SSNR tăng cường theo cách tương ứng theo nét đặc trưng này của tín hiệu audio, và so sánh SSNR tăng cường này với ngưỡng quyết định VAD, nên tỉ lệ phát hiện trượt tín hiệu có hoạt động có thể được giảm.

Ngoài ra, bộ xử lý 801 được tạo cấu hình cụ thể để xác định tín hiệu audio này dưới dạng tín hiệu audio cần được xác định theo SNR dải con của tín hiệu audio này.

Tuỳ ý, theo một phương án, bộ xử lý 801 được tạo cấu hình cụ thể để xác định tín hiệu audio này dưới dạng tín hiệu audio cần được xác định trong trường hợp số lượng dải con đầu cao tần, mà nằm trong tín hiệu audio

này và có các SNR dải con lớn hơn ngưỡng định trước thứ nhất, là lớn hơn số lượng thứ nhất.

Tuỳ ý, theo phương án khác, bộ xử lý 801 được tạo cấu hình cụ thể để xác định tín hiệu audio này dưới dạng tín hiệu audio cần được xác định trong trường hợp số lượng dải con đầu cao tần, mà nằm trong tín hiệu audio này và có các SNR dải con lớn hơn ngưỡng định trước thứ nhất, là lớn hơn số lượng thứ hai, và số lượng dải con đầu hạ tần, mà nằm trong tín hiệu audio này và có các SNR dải con nhỏ hơn ngưỡng định trước thứ hai, là lớn hơn số lượng thứ ba.

Ngưỡng định trước thứ nhất và ngưỡng định trước thứ hai có thể thu được bằng phương pháp tập hợp thống kê theo một lượng lớn mẫu tiếng nói. Cụ thể là, số liệu thống kê về các SNR dải con của các dải con đầu cao tần được tập hợp trong một lượng lớn các mẫu tiếng nói âm vô thanh bao gồm tạp âm nền, và ngưỡng định trước thứ nhất được xác định theo các SNR dải con này, sao cho các SNR dải con của hầu hết trong số các dải con đầu cao tần trong các mẫu âm vô thanh này là lớn hơn ngưỡng định trước thứ nhất này. Tương tự, số liệu thống kê về các SNR dải con của các dải con đầu hạ tần là được tập hợp trong các mẫu tiếng nói âm vô thanh này, và ngưỡng định trước thứ hai được xác định theo các SNR dải con này, sao cho các SNR dải con của hầu hết trong số các dải con đầu hạ tần trong các mẫu tiếng nói âm vô thanh này là nhỏ hơn ngưỡng định trước thứ hai này.

Số lượng thứ nhất, số lượng thứ hai, và số lượng thứ ba cũng được thu thập bằng phương pháp tập hợp thống kê. Số lượng thứ nhất được lấy làm ví dụ, trong đó trong một lượng lớn khung mẫu tiếng nói âm vô thanh bao gồm tạp âm, thì số liệu thống kê về lượng dải con của các dải con đầu cao tần, mà có các SNR dải con lớn hơn ngưỡng định trước thứ nhất, được tập hợp, và số lượng thứ nhất này được xác định theo lượng này, sao cho số lượng dải con đầu cao tần, mà nằm trong hầu hết trong số các khung mẫu tiếng nói âm vô thanh này và có các SNR dải con lớn hơn ngưỡng định trước thứ nhất,

là lớn hơn số lượng thứ nhất. Phương pháp thu thập số lượng thứ hai là tương tự như phương pháp thu thập số lượng thứ nhất. Số lượng thứ hai có thể giống với số lượng thứ nhất, hoặc số lượng thứ hai có thể khác với số lượng thứ nhất. Tương tự, đối với số lượng thứ ba, thì trong một lượng lớn khung mẫu tiếng nói âm vô thanh bao gồm tạp âm, số liệu thống kê về lượng dải con của các dải con đầu hạ tần, mà có các SNR dải con nhỏ hơn ngưỡng định trước thứ hai, được tập hợp, và số lượng thứ ba này được xác định theo lượng này, sao cho số lượng dải con đầu hạ tần, mà nằm trong hầu hết trong số các khung mẫu tiếng nói âm vô thanh này và có các SNR dải con nhỏ hơn ngưỡng định trước thứ hai, là lớn hơn số lượng thứ ba này.

Fig.9 là hình vẽ thể hiện sơ đồ khối cấu trúc của thiết bị khác theo một phương án của sáng chế. Thiết bị 900 trên Fig.9 có thể thực hiện tất cả các bước được thể hiện trên Fig.4. Như được thể hiện trên Fig.9, thiết bị 900 này bao gồm khối xác định thứ nhất 901, khối xác định thứ hai 902, khối xác định thứ ba 903, và khối xác định thứ tư 904.

Khối xác định thứ nhất 901 được tạo cấu hình để xác định tín hiệu audio vào dưới dạng tín hiệu audio cần được xác định.

Khối xác định thứ hai 902 được tạo cấu hình để thu thập SSNR tham chiếu của tín hiệu audio này.

Cụ thể là, SSNR tham chiếu có thể là SSNR thu được sau khi tính toán bằng công thức 1.1.

Khối xác định thứ ba 903 được tạo cấu hình để dùng thuật toán định trước để giảm ngưỡng quyết định VAD tham chiếu, để thu được ngưỡng quyết định VAD đã giảm.

Cụ thể là, ngưỡng quyết định VAD tham chiếu này có thể là ngưỡng quyết định VAD mặc định, và ngưỡng quyết định VAD tham chiếu này có thể được lưu giữ trước, hoặc có thể là tạm thời thu được qua việc tính toán, trong đó ngưỡng quyết định VAD tham chiếu này có thể được tính toán bằng công nghệ đã biết hiện có. Khi ngưỡng quyết định VAD tham chiếu

được giảm bằng thuật toán định trước, thì thuật toán định trước này có thể là nhân ngưỡng quyết định VAD tham chiếu này với hệ số nhỏ hơn 1, hoặc thuật toán khác có thể được sử dụng. Phương án này của sáng chế không áp đặt giới hạn nào đối với thuật toán cụ thể được sử dụng. Ngưỡng quyết định VAD có thể được giảm một cách phù hợp bằng thuật toán định trước, sao cho SSNR tăng cường trở nên lớn hơn ngưỡng quyết định VAD đã giảm. Do đó, xác suất phát hiện trượt tín hiệu có hoạt động có thể được giảm.

Khối xác định thứ tư 904 được tạo cấu hình để so sánh SSNR tham chiếu với ngưỡng quyết định VAD đã giảm để xác định xem tín hiệu audio này có phải là tín hiệu có hoạt động hay không.

Tùy ý, theo một phương án, khối xác định thứ nhất 901 được tạo cấu hình cụ thể để xác định tín hiệu audio này dưới dạng tín hiệu audio cần được xác định theo SNR dải con của tín hiệu audio này.

Tùy ý, theo một phương án, trong trường hợp khối xác định thứ nhất 901 xác định tín hiệu audio này dưới dạng tín hiệu audio cần được xác định theo SNR dải con của tín hiệu audio này, thì khối xác định thứ nhất 901 được tạo cấu hình cụ thể để xác định tín hiệu audio này dưới dạng tín hiệu audio cần được xác định trong trường hợp số lượng dải con đầu cao tần, mà nằm trong tín hiệu audio này và có các SNR dải con lớn hơn ngưỡng định trước thứ nhất, là lớn hơn số lượng thứ nhất.

Tùy ý, theo một phương án, trong trường hợp khối xác định thứ nhất 901 xác định tín hiệu audio này dưới dạng tín hiệu audio cần được xác định theo SNR dải con của tín hiệu audio này, thì khối xác định thứ nhất 901 được tạo cấu hình cụ thể để xác định tín hiệu audio này dưới dạng tín hiệu audio cần được xác định trong trường hợp số lượng dải con đầu cao tần, mà nằm trong tín hiệu audio này và có các SNR dải con lớn hơn ngưỡng định trước thứ nhất, là lớn hơn số lượng thứ hai, và số lượng dải con đầu hạ tần, mà nằm trong tín hiệu audio này và có các SNR dải con nhỏ hơn ngưỡng định trước thứ hai, là lớn hơn số lượng thứ ba.

Tùy ý, theo một phương án, trong trường hợp khối xác định thứ nhất 901 xác định tín hiệu audio này dưới dạng tín hiệu audio cần được xác định theo SNR dải con của tín hiệu audio này, thì khối xác định thứ nhất 901 được tạo cấu hình cụ thể để xác định tín hiệu audio này dưới dạng tín hiệu audio cần được xác định trong trường hợp số lượng dải con, mà nằm trong tín hiệu audio này và có giá trị của các SNR dải con lớn hơn ngưỡng định trước thứ ba, là lớn hơn số lượng thứ tư.

Tùy ý, theo một phương án, khối xác định thứ nhất 901 được tạo cấu hình cụ thể để xác định tín hiệu audio này dưới dạng tín hiệu audio cần được xác định trong trường hợp xác định được rằng tín hiệu audio này là tín hiệu âm vô thanh. Cụ thể là, người có kiến thức trung bình trong lĩnh vực này có thể thấy rằng có thể có nhiều phương pháp để phát hiện xem tín hiệu audio này có phải là tín hiệu âm vô thanh hay không. Ví dụ, việc tín hiệu audio này có phải là tín hiệu âm vô thanh hay không có thể được xác định bằng cách phát hiện ZCR của tín hiệu audio này. Cụ thể là, trong trường hợp mà ZCR của tín hiệu audio này lớn hơn ngưỡng ZCR, thì tín hiệu audio này được xác định là tín hiệu âm vô thanh, trong đó ngưỡng ZCR này được xác định theo một lượng lớn thực nghiệm.

Ngưỡng định trước thứ nhất và ngưỡng định trước thứ hai có thể thu được bằng phương pháp tập hợp thống kê theo một lượng lớn mẫu tiếng nói. Cụ thể là, số liệu thống kê về các SNR dải con của các dải con đầu cao tần được tập hợp trong một lượng lớn các mẫu tiếng nói âm vô thanh bao gồm tạp âm nền, và ngưỡng định trước thứ nhất được xác định theo các SNR dải con này, sao cho các SNR dải con của hầu hết trong số các dải con đầu cao tần trong các mẫu âm vô thanh này là lớn hơn ngưỡng định trước thứ nhất này. Tương tự, số liệu thống kê về các SNR dải con của các dải con đầu hạ tần là được tập hợp trong các mẫu tiếng nói âm vô thanh này, và ngưỡng định trước thứ hai được xác định theo các SNR dải con này, sao cho các SNR dải con của hầu hết trong số các dải con đầu hạ tần trong các mẫu tiếng

nói âm vô thanh này là nhỏ hơn ngưỡng định trước thứ hai này.

Ngưỡng định trước thứ ba cũng được thu thập bằng phương pháp tập hợp thống kê. Cụ thể là, ngưỡng định trước thứ ba được xác định theo các SNR dải con của một lượng lớn tín hiệu tạp âm, sao cho các SNR dải con của hầu hết trong số các dải con trong các tín hiệu tạp âm này là nhỏ hơn ngưỡng định trước thứ ba này.

Số lượng thứ nhất, số lượng thứ hai, số lượng thứ ba, và số lượng thứ tư cũng được thu thập bằng phương pháp tập hợp thống kê. Số lượng thứ nhất được lấy làm ví dụ, trong đó trong một lượng lớn mẫu tiếng nói bao gồm tạp âm, thì số liệu thống kê về lượng dải con của các dải con đầu cao tần, mà có các SNR dải con lớn hơn ngưỡng định trước thứ nhất, được tập hợp, và số lượng thứ nhất này được xác định theo lượng này, sao cho số lượng dải con đầu cao tần, mà nằm trong hầu hết trong số các mẫu tiếng nói này và có các SNR dải con lớn hơn ngưỡng định trước thứ nhất, là lớn hơn số lượng thứ nhất. Phương pháp để xác định số lượng thứ hai là tương tự như phương pháp để xác định số lượng thứ nhất. Số lượng thứ hai có thể giống với số lượng thứ nhất, hoặc có thể khác với số lượng thứ nhất. Tương tự, đối với số lượng thứ ba, thì trong một lượng lớn mẫu tiếng nói bao gồm tạp âm, số liệu thống kê về lượng dải con của các dải con đầu hạ tần, mà có các SNR dải con lớn hơn ngưỡng định trước thứ hai, được tập hợp, và số lượng thứ ba này được xác định theo lượng này, sao cho số lượng dải con đầu hạ tần, mà nằm trong hầu hết trong số các mẫu tiếng nói này và có các SNR dải con lớn hơn ngưỡng định trước thứ hai, là lớn hơn số lượng thứ ba này. Đối với số lượng thứ tư, trong lượng lớn mẫu tiếng nói bao gồm tạp âm này, thì số liệu thống kê về số lượng dải con, mà có các SNR dải con lớn hơn ngưỡng định trước thứ ba, được tập hợp, và số lượng thứ tư này được xác định theo lượng này, sao cho số lượng dải con, mà nằm trong hầu hết trong số các mẫu tiếng nói này và có các SNR dải con lớn hơn ngưỡng định trước thứ ba, là lớn hơn số lượng thứ tư này.

Thiết bị 900 trên Fig.9 có thể xác định nét đặc trưng của tín hiệu audio vào, giảm ngưỡng quyết định VAD tham chiếu theo nét đặc trưng của tín hiệu audio này, và so sánh SSNR tăng cường với ngưỡng quyết định VAD đã giảm, nên có thể giảm được tỉ lệ phát hiện trượt tín hiệu có hoạt động.

Fig.10 là hình vẽ thể hiện sơ đồ khối cấu trúc của thiết bị khác theo một phương án của sáng chế. Thiết bị 1000 trên Fig.10 có thể thực hiện tất cả các bước được thể hiện trên Fig.4. Như được thể hiện trên Fig.10, thiết bị 1000 này bao gồm bộ xử lý 1001 và bộ nhớ 1002. Bộ xử lý 1001 có thể là bộ xử lý thông dụng, DSP, ASIC, FPGA hoặc thành phần logic lập trình được khác, cổng rời rạc hoặc thành phần logic tranzito, hoặc thành phần phần cứng rời rạc, mà có thể thực hiện các phương pháp, các bước, và các sơ đồ khối logic đã được bộc lộ theo các phương án của sáng chế. Bộ xử lý thông dụng có thể là bộ vi xử lý, hoặc bộ xử lý nêu trên có thể là bộ xử lý thông thường bất kì hoặc các phương tiện tương tự. Các bước của các phương pháp đã được bộc lộ theo các phương án của sáng chế có thể được thực hiện trực tiếp bởi bộ xử lý giải mã phần cứng, hoặc được thực hiện bởi tổ hợp của các môđun phần cứng và phần mềm trong bộ xử lý giải mã. Môđun phần mềm này có thể được đặt trong phương tiện lưu trữ đã biết trong lĩnh vực, chẳng hạn RAM, bộ nhớ flash, ROM, bộ nhớ chỉ đọc lập trình được, bộ nhớ lập trình và xoá được bằng điện, hoặc thanh ghi. Phương tiện lưu trữ này được đặt trong bộ nhớ 1002. Bộ xử lý 1001 đọc lệnh từ bộ nhớ 1002, và thực hiện các bước của các phương pháp nêu trên cùng với phần cứng.

Bộ xử lý 1001 được tạo cấu hình để xác định tín hiệu audio vào dưới dạng tín hiệu audio cần được xác định.

Bộ xử lý 1001 được tạo cấu hình để thu thập SSNR tham chiếu của tín hiệu audio này.

Cụ thể là, SSNR tham chiếu có thể là SSNR thu được sau khi tính toán bằng công thức 1.1.

Bộ xử lý 1001 được tạo cấu hình để dùng thuật toán định trước để giảm ngưỡng quyết định VAD tham chiếu, để thu được ngưỡng quyết định VAD đã giảm.

Cụ thể là, ngưỡng quyết định VAD tham chiếu này có thể là ngưỡng quyết định VAD mặc định, và ngưỡng quyết định VAD tham chiếu này có thể được lưu giữ trước, hoặc có thể là tạm thời thu được qua việc tính toán, trong đó ngưỡng quyết định VAD tham chiếu này có thể được tính toán bằng công nghệ đã biết hiện có. Khi ngưỡng quyết định VAD tham chiếu được giảm bằng thuật toán định trước, thì thuật toán định trước này có thể là nhân ngưỡng quyết định VAD tham chiếu này với hệ số nhỏ hơn 1, hoặc thuật toán khác có thể được sử dụng. Phương án này của sáng chế không áp đặt giới hạn nào đối với thuật toán cụ thể được sử dụng. Ngưỡng quyết định VAD có thể được giảm một cách phù hợp bằng thuật toán định trước, sao cho SSNR tăng cường trở nên lớn hơn ngưỡng quyết định VAD đã giảm. Do đó, xác suất phát hiện trượt tín hiệu có hoạt động có thể được giảm.

Bộ xử lý 1001 được tạo cấu hình để so sánh SSNR tham chiếu với ngưỡng quyết định VAD đã giảm để xác định xem tín hiệu audio này có phải là tín hiệu có hoạt động hay không.

Tùy ý, theo một phương án, bộ xử lý 1001 được tạo cấu hình cụ thể để xác định tín hiệu audio này dưới dạng tín hiệu audio cần được xác định theo SNR dải con của tín hiệu audio này.

Tùy ý, theo một phương án, trong trường hợp bộ xử lý 1001 xác định tín hiệu audio này dưới dạng tín hiệu audio cần được xác định theo SNR dải con của tín hiệu audio này, thì bộ xử lý 1001 được tạo cấu hình cụ thể để xác định tín hiệu audio này dưới dạng tín hiệu audio cần được xác định trong trường hợp số lượng dải con đầu cao tần, mà nằm trong tín hiệu audio này và có các SNR dải con lớn hơn ngưỡng định trước thứ nhất, là lớn hơn số lượng thứ nhất.

Tùy ý, theo một phương án, trong trường hợp bộ xử lý 1001 xác định

tín hiệu audio này dưới dạng tín hiệu audio cần được xác định theo SNR dải con của tín hiệu audio này, thì bộ xử lý 1001 được tạo cấu hình cụ thể để xác định tín hiệu audio này dưới dạng tín hiệu audio cần được xác định trong trường hợp số lượng dải con đầu cao tần, mà nằm trong tín hiệu audio này và có các SNR dải con lớn hơn ngưỡng định trước thứ nhất, là lớn hơn số lượng thứ hai, và số lượng dải con đầu hạ tần, mà nằm trong tín hiệu audio này và có các SNR dải con nhỏ hơn ngưỡng định trước thứ hai, là lớn hơn số lượng thứ ba.

Tuỳ ý, theo một phương án, trong trường hợp bộ xử lý 1001 xác định tín hiệu audio này dưới dạng tín hiệu audio cần được xác định theo SNR dải con của tín hiệu audio này, thì bộ xử lý 1001 được tạo cấu hình cụ thể để xác định tín hiệu audio này dưới dạng tín hiệu audio cần được xác định trong trường hợp số lượng dải con, mà nằm trong tín hiệu audio này và có giá trị của các SNR dải con lớn hơn ngưỡng định trước thứ ba, là lớn hơn số lượng thứ tư.

Tuỳ ý, theo một phương án, bộ xử lý 1001 được tạo cấu hình cụ thể để xác định tín hiệu audio này dưới dạng tín hiệu audio cần được xác định trong trường hợp xác định được rằng tín hiệu audio này là tín hiệu âm vô thanh. Cụ thể là, người có kiến thức trung bình trong lĩnh vực này có thể thấy rằng có thể có nhiều phương pháp để phát hiện xem tín hiệu audio này có phải là tín hiệu âm vô thanh hay không. Ví dụ, việc tín hiệu audio này có phải là tín hiệu âm vô thanh hay không có thể được xác định bằng cách phát hiện ZCR của tín hiệu audio này. Cụ thể là, trong trường hợp mà ZCR của tín hiệu audio này lớn hơn ngưỡng ZCR, thì tín hiệu audio này được xác định là tín hiệu âm vô thanh, trong đó ngưỡng ZCR này được xác định theo một lượng lớn thực nghiệm.

Ngưỡng định trước thứ nhất và ngưỡng định trước thứ hai có thể thu được bằng phương pháp tập hợp thống kê theo một lượng lớn mẫu tiếng nói. Cụ thể là, số liệu thống kê về các SNR dải con của các dải con đầu cao tần

được tập hợp trong một lượng lớn các mẫu tiếng nói âm vô thanh bao gồm tạp âm nền, và ngưỡng định trước thứ nhất được xác định theo các SNR dải con này, sao cho các SNR dải con của hầu hết trong số các dải con đầu cao tần trong các mẫu âm vô thanh này là lớn hơn ngưỡng định trước thứ nhất này. Tương tự, số liệu thống kê về các SNR dải con của các dải con đầu hạ tần là được tập hợp trong các mẫu tiếng nói âm vô thanh này, và ngưỡng định trước thứ hai được xác định theo các SNR dải con này, sao cho các SNR dải con của hầu hết trong số các dải con đầu hạ tần trong các mẫu tiếng nói âm vô thanh này là nhỏ hơn ngưỡng định trước thứ hai này.

Ngưỡng định trước thứ ba cũng được thu thập bằng phương pháp tập hợp thống kê. Cụ thể là, ngưỡng định trước thứ ba được xác định theo các SNR dải con của một lượng lớn tín hiệu tạp âm, sao cho các SNR dải con của hầu hết trong số các dải con trong các tín hiệu tạp âm này là nhỏ hơn ngưỡng định trước thứ ba này.

Số lượng thứ nhất, số lượng thứ hai, số lượng thứ ba, và số lượng thứ tư cũng được thu thập bằng phương pháp tập hợp thống kê. Số lượng thứ nhất được lấy làm ví dụ, trong đó trong một lượng lớn mẫu tiếng nói bao gồm tạp âm, thì số liệu thống kê về lượng dải con của các dải con đầu cao tần, mà có các SNR dải con lớn hơn ngưỡng định trước thứ nhất, được tập hợp, và số lượng thứ nhất này được xác định theo lượng này, sao cho số lượng dải con đầu cao tần, mà nằm trong hầu hết trong số các mẫu tiếng nói này và có các SNR dải con lớn hơn ngưỡng định trước thứ nhất, là lớn hơn số lượng thứ nhất. Phương pháp để xác định số lượng thứ hai là tương tự như phương pháp để xác định số lượng thứ nhất. Số lượng thứ hai có thể giống với số lượng thứ nhất, hoặc có thể khác với số lượng thứ nhất. Tương tự, đối với số lượng thứ ba, thì trong một lượng lớn mẫu tiếng nói bao gồm tạp âm, số liệu thống kê về lượng dải con của các dải con đầu hạ tần, mà có các SNR dải con lớn hơn ngưỡng định trước thứ hai, được tập hợp, và số lượng thứ ba này được xác định theo lượng này, sao cho số lượng dải con

đầu hạ tần, mà nằm trong hầu hết trong số các mẫu tiếng nói này và có các SNR dải con lớn hơn ngưỡng định trước thứ hai, là lớn hơn số lượng thứ ba này. Đối với số lượng thứ tư, trong lượng lớn mẫu tiếng nói bao gồm tập âm này, thì số liệu thống kê về số lượng dải con, mà có các SNR dải con lớn hơn ngưỡng định trước thứ ba, được tập hợp, và số lượng thứ tư này được xác định theo lượng này, sao cho số lượng dải con, mà nằm trong hầu hết trong số các mẫu tiếng nói này và có các SNR dải con lớn hơn ngưỡng định trước thứ ba, là lớn hơn số lượng thứ tư này.

Thiết bị 1000 trên Fig.10 có thể xác định nét đặc trưng của tín hiệu audio vào, giảm ngưỡng quyết định VAD tham chiếu theo nét đặc trưng của tín hiệu audio này, và so sánh SSNR tăng cường với ngưỡng quyết định VAD đã giảm, nên có thể giảm được tỉ lệ phát hiện trượt tín hiệu có hoạt động.

Dựa vào các ví dụ được mô tả trong các phương án trong phần mô tả này, người có kiến thức trung bình trong lĩnh vực này có thể thấy rằng các khối và các bước thuật toán nêu trên có thể được thực hiện bằng phần cứng điện tử, hoặc tổ hợp giữa phần mềm máy tính và phần cứng điện tử. Việc các chức năng này được thực hiện bằng phần cứng hay phần mềm phụ thuộc vào các ứng dụng cụ thể và các điều kiện ràng buộc về thiết kế kỹ thuật. Người có kiến thức trung bình trong lĩnh vực này có thể sử dụng các phương pháp khác nhau để thực hiện các chức năng được mô tả đối với mỗi ứng dụng cụ thể, nhưng điều này không có nghĩa là cách thức thực hiện này nằm ngoài phạm vi của sáng chế.

Người có kiến thức trung bình trong lĩnh vực này có thể thấy rõ rằng, để tiện lợi cho việc mô tả và nhằm mục đích mô tả vắn tắt, thì quá trình hoạt động chi tiết của hệ thống, thiết bị, và các đơn vị nêu trên có thể được tìm thấy ở quá trình tương ứng trong các phương án về phương pháp trên đây, nên không được mô tả lại nữa.

Theo một số phương án trong đơn này, cần hiểu rằng hệ thống, thiết bị

và phương pháp được bộc lộ có thể được thực hiện theo những cách khác. Ví dụ, phương án về thiết bị đã được mô tả là chỉ được nêu làm ví dụ. Ví dụ, nhóm đơn vị nêu trên chỉ là nhóm chức năng logic, và nó có thể là nhóm khác khi thực hiện thực tế. Ví dụ, các đơn vị hoặc các thành phần có thể được kết hợp hoặc được tích hợp vào hệ thống khác, hoặc một số dấu hiệu có thể được bỏ qua, hoặc không được thực hiện. Ngoài ra, những sự kết nối với nhau hoặc những sự kết nối hoặc các kết nối giao tiếp trực tiếp đã được thể hiện hoặc được mô tả nêu trên là có thể được thực hiện nhờ sử dụng một số giao diện. Các mối ghép hoặc các kết nối giao tiếp gián tiếp giữa các thiết bị hoặc các khối là có thể được thực hiện về mặt điện tử, cơ học, hoặc các dạng khác.

Các đơn vị được mô tả dưới dạng các bộ phận riêng rẽ có thể là, hoặc không phải là, riêng rẽ về mặt vật lý, và các bộ phận được thể hiện dưới dạng các đơn vị có thể là, hoặc không phải là, các đơn vị vật lý, có thể được đặt tại một vị trí, hoặc có thể được rải rác trên nhiều đơn vị mạng. Một số hoặc tất cả trong số các đơn vị này có thể được chọn theo các nhu cầu thực tế để đạt được các mục đích của các phương án này.

Ngoài ra, các đơn vị chức năng ở các phương án của sáng chế có thể được tích hợp vào một bộ xử lý, hoặc mỗi trong số các đơn vị này có thể tồn tại một mình về mặt vật lý, hoặc hai hoặc nhiều đơn vị được hợp nhất thành một đơn vị.

Khi các chức năng này được thực hiện dưới dạng đơn vị chức năng phần mềm và được bán hoặc được sử dụng dưới dạng sản phẩm độc lập, thì các chức năng này có thể được lưu giữ trên phương tiện lưu trữ đọc được bằng máy tính. Do đó, giải pháp của sáng chế, hoặc phần khắc phục nhược điểm của giải pháp đã biết, hoặc một phần của các giải pháp kỹ thuật này, có thể được thực hiện dưới dạng sản phẩm phần mềm. Sản phẩm phần mềm này được lưu giữ trên phương tiện lưu trữ, và bao gồm một số lệnh để ra lệnh cho thiết bị máy tính (có thể là máy tính cá nhân, máy chủ, hoặc thiết bị

mạng) hoặc bộ xử lý thực hiện toàn bộ hoặc một phần của các bước của các phương pháp đã được mô tả trong các phương án theo sáng chế. Phương tiện lưu trữ nêu trên bao gồm: các phương tiện bất kì mà có thể lưu giữ mã chương trình, chẳng hạn ổ đĩa USB (Universal Serial Bus - buýt nối tiếp vạn năng), đĩa cứng tháo ra được, ROM, RAM, đĩa từ, hoặc đĩa quang.

Phần mô tả nêu trên chỉ nêu các phương án cụ thể của sáng chế, chứ không nhằm giới hạn phạm vi bảo hộ của sáng chế. Các phương án biến thể hoặc thay thế bất kì mà người có kiến thức trung bình trong lĩnh vực này tạo ra trong phạm vi kỹ thuật của sáng chế cũng đều nằm trong phạm vi bảo hộ của sáng chế. Do đó, phạm vi bảo hộ của sáng chế được xác định theo phạm vi bảo hộ của các điểm yêu cầu bảo hộ kèm theo.

YÊU CẦU BẢO HỘ

1. Phương pháp phát hiện tín hiệu audio, trong đó phương pháp này bao gồm các bước:

xác định tín hiệu audio này dưới dạng tín hiệu audio cần được xác định nếu số lượng dải con đầu cao tần, mà nằm trong tín hiệu audio này và có các SNR dải con lớn hơn ngưỡng định trước thứ nhất, là lớn hơn số lượng thứ nhất;

xác định tỉ số tín hiệu trên tạp âm theo phân đoạn (Segmental Signal-to-Noise Ratio - SSNR) tăng cường của tín hiệu audio này, trong đó SSNR tăng cường này lớn hơn SSNR tham chiếu; và

so sánh SSNR tăng cường này với ngưỡng quyết định phát hiện hoạt động thoại (Voice Activity Detection - VAD) để xác định xem tín hiệu audio này có phải là tín hiệu có hoạt động hay không,

trong đó bước xác định SSNR tăng cường của tín hiệu audio là các bước:

xác định SSNR tham chiếu của tín hiệu audio này; và

xác định SSNR tăng cường theo SSNR tham chiếu của tín hiệu audio này.

2. Phương pháp theo điểm 1, trong đó bước xác định SSNR tăng cường theo SSNR tham chiếu của tín hiệu audio là bước:

xác định SSNR tăng cường bằng công thức sau:

$$SSNR' = x * SSNR + y, \text{ trong đó}$$

SSNR biểu thị SSNR tham chiếu, $SSNR'$ biểu thị SSNR tăng cường, và x và y biểu thị các thông số tăng cường.

3. Phương pháp theo điểm 1, trong đó bước xác định SSNR tăng cường

bằng công thức sau:

$$SSNR' = f(x) * SSNR + h(y),$$

trong đó SSNR biểu thị SSNR tham chiếu, SSNR' biểu thị SSNR tăng cường, và $f(x)$ và $h(y)$ biểu thị các hàm tăng cường, và $h(y)$ là hàm liên quan đến NR dài hạn của tín hiệu âm thanh.

4. Phương pháp theo điểm bất kỳ trong số các điểm từ 1 đến 3, trong đó tín hiệu âm thanh đầu vào được phân chia thành N dải con, $N=20$, N dải con bao gồm dải con 0 đến dải con 19, và dải con 18 và dải con 19 là các dải con đầu cao tần.

5. Phương pháp phát hiện tín hiệu audio, trong đó phương pháp này bao gồm các bước:

xác định tín hiệu audio vào dưới dạng tín hiệu audio cần được xác định;

xác định trọng số của tỉ số tín hiệu trên tạp âm (SNR) theo dải con của mỗi dải con trong tín hiệu audio này, trong đó trọng số của SNR dải con của dải con đầu cao tần mà có SNR dải con lớn hơn ngưỡng định trước thứ nhất là lớn hơn trọng số của SNR dải con của dải con khác;

xác định tỉ số tín hiệu trên tạp âm theo phân đoạn (SSNR) tăng cường theo SNR dải con của mỗi dải con và trọng số của SNR dải con của mỗi dải con trong tín hiệu audio này, trong đó SSNR tăng cường này là lớn hơn SSNR tham chiếu; và

so sánh SSNR tăng cường này với ngưỡng quyết định phát hiện hoạt động thoại (Voice Activity Detection - VAD) để xác định xem tín hiệu audio này có phải là tín hiệu có hoạt động hay không.

6. Phương pháp theo điểm 5, trong đó tín hiệu âm thanh đầu vào được phân chia thành N dải con, $N=20$, N dải con bao gồm dải con 0 đến dải con 19, và

dải con 18 và dải con 19 là các dải con đầu cao tần.

7. Thiết bị phát hiện tín hiệu audio, trong đó thiết bị này bao gồm:

khối xác định thứ nhất, được tạo cấu hình để xác định tín hiệu audio vào dưới dạng tín hiệu audio cần được xác định;

khối xác định thứ hai, được tạo cấu hình để xác định tỉ số tín hiệu trên tạp âm theo phân đoạn (SSNR) tăng cường của tín hiệu audio này, trong đó SSNR tăng cường này lớn hơn SSNR tham chiếu; và

khối xác định thứ ba, được tạo cấu hình để so sánh SSNR tăng cường này với ngưỡng quyết định phát hiện hoạt động thoại (VAD) để xác định xem tín hiệu audio này có phải là tín hiệu có hoạt động hay không,

trong đó khối xác định thứ hai được tạo cấu hình để xác định SSNR tham chiếu của tín hiệu âm thanh và xác định SSNR được tăng cường theo SSNR tham chiếu của tín hiệu âm thanh.

8. Thiết bị theo điểm 7, trong đó khối xác định thứ hai được tạo cấu hình để xác định SSNR tăng cường bằng công thức sau:

$$SSNR' = x * SSNR + y, \text{ trong đó}$$

SSNR biểu thị SSNR tham chiếu, $SSNR'$ biểu thị SSNR tăng cường, và x và y biểu thị các thông số tăng cường.

9. Phương pháp theo điểm 7, trong đó bước xác định SSNR tăng cường bằng công thức sau:

$$SSNR' = f(x) * SSNR + h(y),$$

trong đó SSNR biểu thị SSNR tham chiếu, $SSNR'$ biểu thị SSNR tăng cường, và $f(x)$ và $h(y)$ biểu thị các hàm tăng cường, và $h(y)$ là hàm liên quan đến NR dài hạn của tín hiệu âm thanh.

10. Phương pháp theo điểm bất kỳ trong số các điểm từ 7 đến 9, trong đó tín hiệu âm thanh đầu vào được phân chia thành N dải con, $N=20$, N dải con bao gồm dải con 0 đến dải con 19, và dải con 18 và dải con 19 là các dải con đầu cao tần.

11. Thiết bị phát hiện tín hiệu audio, trong đó thiết bị này bao gồm:

khối xác định thứ nhất, được tạo cấu hình để xác định tín hiệu audio vào dưới dạng tín hiệu audio cần được xác định;

khối xác định thứ hai, được tạo cấu hình để xác định trọng số của tỉ số tín hiệu trên tạp âm (SNR) theo dải con của mỗi dải con trong tín hiệu audio này, trong đó trọng số của SNR dải con của dải con đầu cao tần, mà có SNR dải con lớn hơn ngưỡng định trước thứ nhất, là lớn hơn trọng số của SNR dải con của dải con khác, và xác định tỉ số tín hiệu trên tạp âm theo phân đoạn (SSNR) tăng cường theo SNR dải con của mỗi dải con và trọng số của SNR dải con của mỗi dải con trong tín hiệu audio này, trong đó SSNR tăng cường này là lớn hơn SSNR tham chiếu; và

khối xác định thứ ba, được tạo cấu hình để so sánh SSNR tăng cường này với ngưỡng quyết định phát hiện hoạt động thoại (VAD) để xác định xem tín hiệu audio này có phải là tín hiệu có hoạt động hay không.

12. Phương pháp theo điểm 11, trong đó tín hiệu âm thanh đầu vào được phân chia thành N dải con, $N=20$, N dải con bao gồm dải con 0 đến dải con 19, và dải con 18 và dải con 19 là các dải con đầu cao tần.

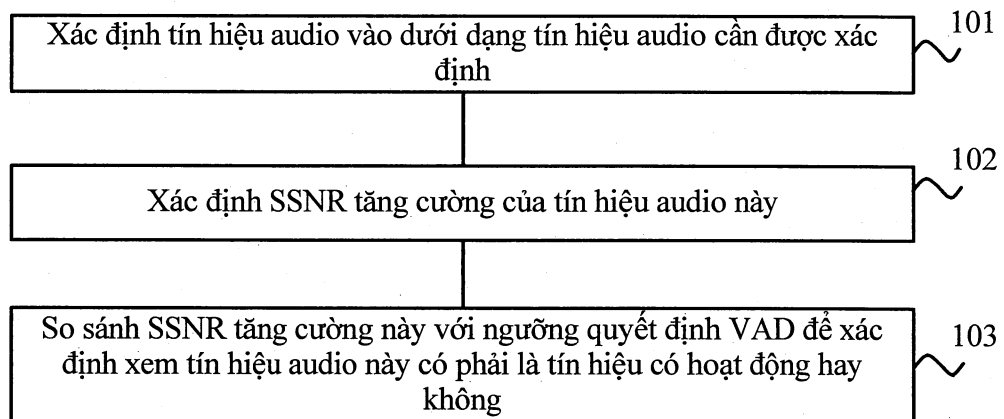


Fig.1

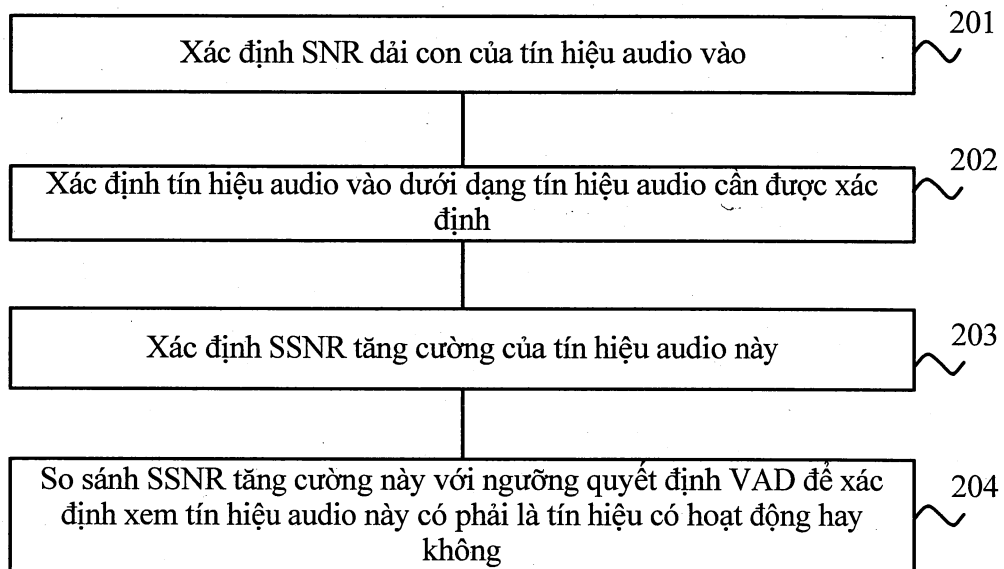


Fig.2

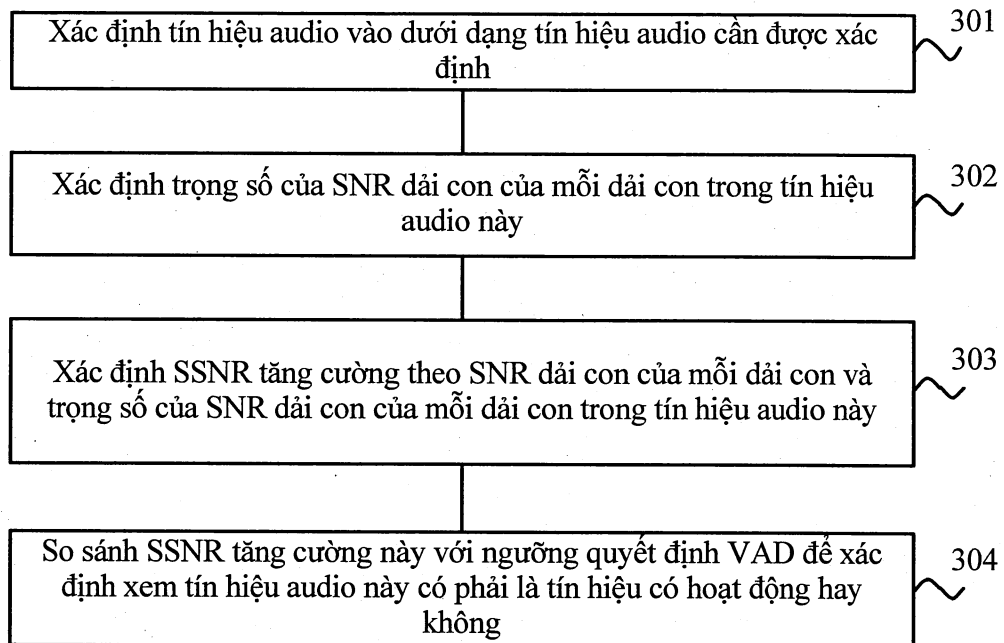


Fig.3

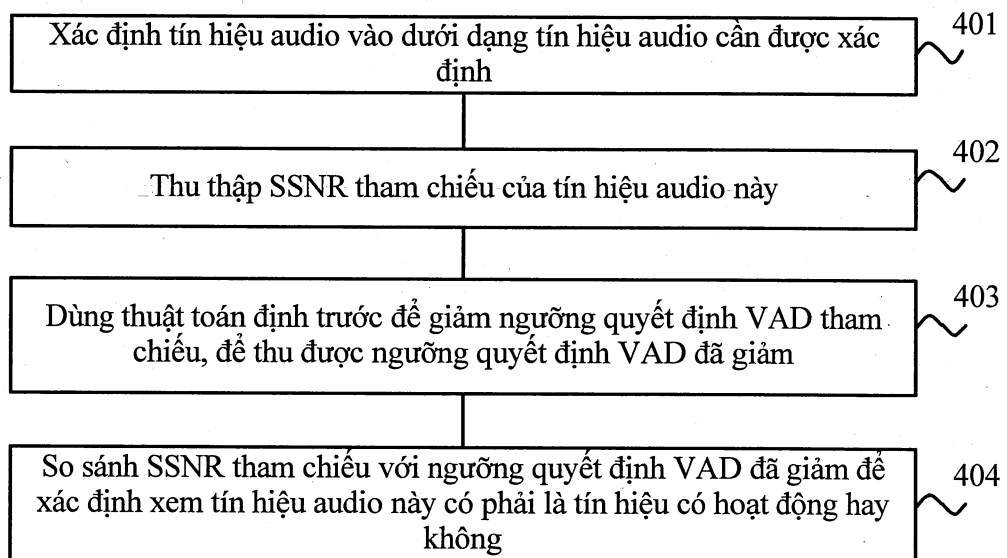


Fig.4

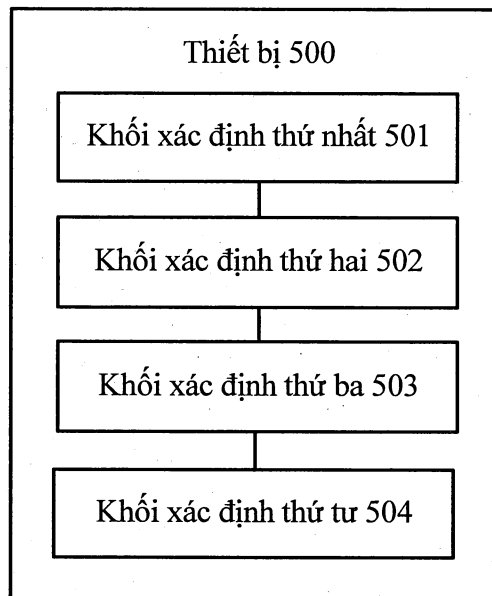


Fig.5

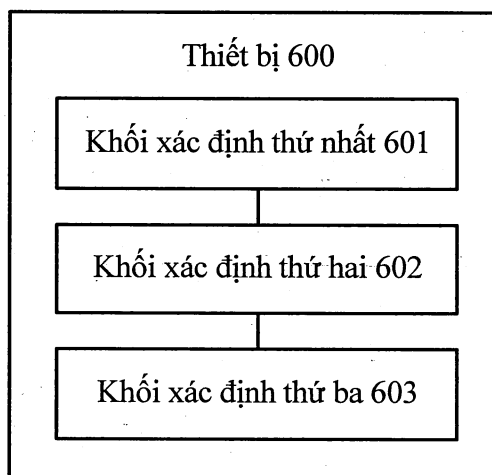


Fig.6

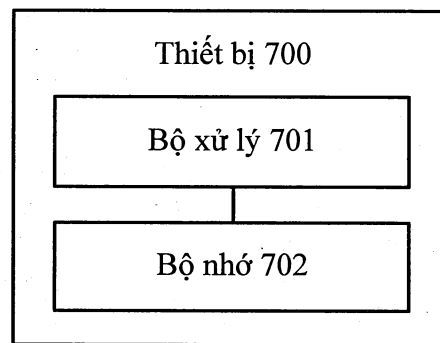


Fig.7

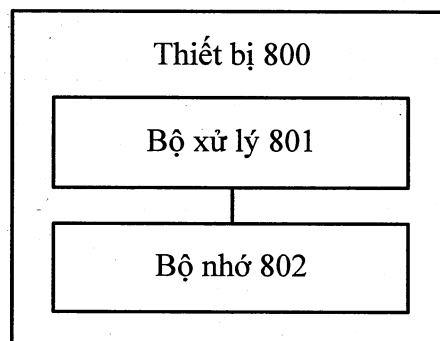


Fig.8

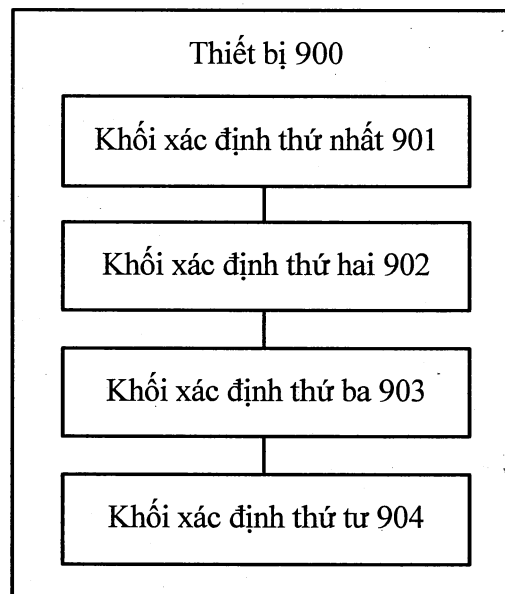


Fig.9

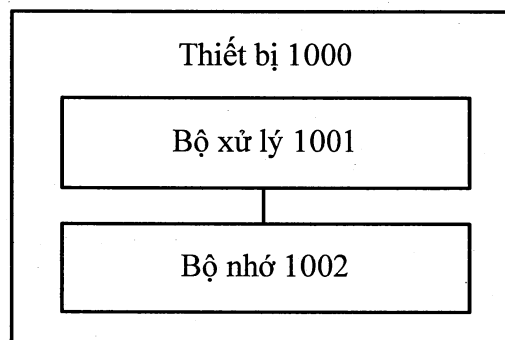


Fig.10